Curating Local Knowledge

Dalton, Patricio; Rüschenpöhler, Julius; Uras, Burak; Zia, Bilal

Publication date:
2020

Document Version
Early version, also known as pre-print

No. 2020-033

CURATING LOCAL KNOWLEDGE: EXPERIMENTAL EVIDENCE FROM SMALL RETAILERS IN INDONESIA

By

Patricio S. Dalton, Julius Rüschenpöhler,
Burak Uras, Bilal Zia

Revision of CentER DP 2019-015

15 December 2020

ISSN 0924-7815
ISSN 2213-9532

Curating Local Knowledge: Experimental Evidence from Small Retailers in Indonesia*

Patricio S. Dalton† Julius Rüschenpöhler‡ Burak Uras† Bilal Zia§

December 14, 2020

Abstract

Business practices and performance vary widely across businesses within the same sector. A key outstanding question is why profitable practices do not readily diffuse. We conduct a field experiment among urban retailers in Indonesia to study whether alleviating informational and behavioral frictions can facilitate such diffusion in a cost-effective manner. Through quantitative and qualitative fieldwork, we curate a handbook that associates locally relevant practices with performance, and provides idiosyncratic implementation guidance informed by exemplary local retailers. We complement this handbook with two light-touch interventions to facilitate behavior change. A subset of retailers is invited to a documentary movie screening featuring the paths to success of exemplary peers. Another subset is offered two 30 minute personal visits by a local facilitator. A third group is offered both. Eighteen months later, we find significant impacts on practice adoption when the handbook is coupled with the two behavioral nudges, and up to a 35% increase in profits and 16.7% increase in sales. These findings suggest both informational and behavioral constraints are at play. The types of practices adopted map the performance improvements to efficiency gains rather than other channels. A simple cost-benefit analysis shows such locally relevant knowledge can be codified and scaled successfully at relatively low cost.

Keywords: Business Growth; Efficiency Gains; Small-scale Enterprises; Peer Knowledge; Self-Learning; Social Learning.
JEL Codes: O12; L26; M20; O31; O33; O35; O17; M50

*We thank the Abdul Latif Jameel Poverty Action Lab (J-PAL) for hosting our study, in particular Ni Luh Putu Satyaning Pradnya Paramita, Dwitri Amalia, Raisa Annisa, Lukman Edwindra and Gabriel Halm for excellent research assistance. We also thank participants at the 8th Wageningen-Lucerne-Tilburg Development Economics workshop, IPA 8th SME Working Group Meeting at MIT Sloan, 9th Tilburg-Wageningen Development Economics Workshop, ENTER-Jamboree 2018 in Toulouse School of Economics, UCPH Field Days Conference at U. of Copenhagen, NBER 2018 Entrepreneurship Working Group in Boston, SEU 2018 at U. de Montevideo, Innovation Growth Lab 2019 Global Conference in Berlin, the Empirical Management Conference 2019 at The World Bank, The Econometric Society World Congress 2020 at Bocconi; and Economics Seminars at EBRD, Radbout Nijmegen, UC Louvain, Stanford, Tilburg, U. Torcuato Di Tella, U. of Glasgow, U. of Gottingen, U. of Florence, Tinbergen Institute, Erasmus University and the Development Economics Network Berlin. Special thanks to Imran Rasul (Editor), four anonymous referees, Thorsten Beck, Nick Bloom, Erwin Bulte, Michelle Brock, Esther Duflo, David McKenzie, Karen Macours, Rossa O'Keeffe-O'Donovan, Menno Pradhan, Sarada, Antoinette Schoar, Daan van Soest, Ben Vollaard and Chris Woodruff for helpful discussions and comments. This paper was produced under the framework of the "Enabling Innovation and Productivity Growth in Low Income Countries (EIP-LIC/PO5639)" project, funded by the Department for International Development (DFID) of the United Kingdom and implemented by Tilburg University. Additional funding was received from the World Bank Strategic Research Program (SRP) and the Tilburg Economics Department. Research on the ground was conducted in collaboration with J-PAL South-East Asia, SurveyMETER and Micra. AEA RCT Registry ID: 0001175.

†Tilburg University, Department of Economics and CentER, Warandelaan 2, 5037 AB, Tilburg, The Netherlands. E-mail corresponding author: p.s.dalton@uvt.nl (corresponding author), and r.b.uras@uvt.nl

‡Center for Effective Global Action, UC Berkeley. E-mail: ruschenpohler@berkeley.edu.

§The World Bank, Development Research Group, 1818 H St. N.W. Washington, D.C. 20433. E-mail: bzia@worldbank.org.

I Introduction

A large body of empirical evidence on firms shows that there is remarkable heterogeneity in the adoption of business practices within economic sectors (e.g. Bloom and Van Reenen, 2007; Syverson, 2011; Gibbons and Henderson, 2013). Importantly, such heterogeneity is strongly associated with firm-level productivity differences (de Mel et al., 2009; McKenzie and Woodruff, 2017; Bloom et al., 2019). These two stylized facts imply that knowledge on locally profitable business practices does exist, yet it does not diffuse naturally. Understanding these constraints as well as factors and processes that can facilitate knowledge diffusion is therefore an important and outstanding research question. This paper addresses this area of research through a field experiment with small-scale urban retailers in Indonesia.

We hypothesize that small-scale firms in developing countries likely face multiple and potentially overlapping constraints to the adoption of profitable business practices. Specifically, we characterize and distinguish two types of constraints: informational and behavioral, and implement a field experiment to study whether alleviating these constraints can facilitate the diffusion of business practices in a cost-effective manner.[1]

Addressing informational constraints is challenging given the muted impacts of standard business training programs found in the literature that ostensibly are designed to target informational asymmetries (Quinn and Woodruff, 2019; McKenzie, 2020). For this reason, we take a fundamentally different approach than standard business training. Instead of providing set-courses, we harness the existing local heterogeneity observed in practices and profits and curate a handbook that associates practical and locally relevant best practices with performance. This handbook identifies and prioritizes particular types of information that are most relevant for performance improvements in the local setting, and provides idiosyncratic implementation guidance informed by exemplary local retailers.

We complement the handbook with two light-touch interventions to address potentially complementary behavioral constraints. A subset of retailers is invited to a movie screening featuring the paths to success of exemplary peers. Another subset is offered two 30 minute in-person counseling visits by a trained facilitator. A third group is offered both. These two behavioral nudges, the movie and counseling, are grounded in the hypothesis that information alone, even if curated, may not be sufficient to facilitate sustained behavior change. The movie targets potential aspirations constraints (Dalton et al., 2016), while the in-person counseling targets constraints related to inattention (Steiner et al., 2017), status-quo bias (Kahneman et al., 1991), and procrastination (Duflo et al., 2011). The combined treatment of both the movie and counseling offers potentially complementary benefits.

Our experiment design consists of 1,301 urban retail shop owners randomized into four treatment groups of 260 retailers and a *Control* group (N=261). All 1040 treated retailers are offered a free copy of the handbook. While retailers assigned to the *Handbook* group receive the handbook alone, those assigned to the *Movie* group additionally receive an invitation to the movie screening, and those assigned to the *Counseling* group additionally receive an invitation to the counseling visits. Finally, retailers assigned to the *All Three* group receive the handbook and the two invitations.

The analysis presented in this paper is based on pooled intent-to-treat (ITT) estimates from

[1]The term "behavioral constraints" encompasses psychological biases (e.g. status-quo, present bias, aspirations) and cognitive shortcuts (e.g. inattention) that can hinder behavior change.

two follow-up surveys, six and eighteen months after the intervention. First, we find underwhelming evidence for the effectiveness of the handbook alone and find no significant improvements in the adoption of business practices. By contrast, we find strong and statistically significant impacts when the handbook is complemented with the behavioral interventions. We find that adding either the movie, counseling, or both result in sizable and statistically significant improvements over *Control* on an aggregate score of all practices covered in the handbook. Specifically, our results show a 12.5% improvement (0.22 s.d.) for *Movie*, a 10.1% improvement (0.18 s.d.) for *Counseling*, and a 16.0% improvement (0.29 s.d.) for *All Three*. These coefficients also reflect statistically significant improvements over providing the handbook alone. Unpacking the aggregate effects into individual practice components, we find statistically significant improvements in all five dimensions of practices for the *All Three* group: record keeping (27.5%), planning (16.9%), stocking up (11.3%), marketing (24.4%), and joint decision making (21.9%). These results are robust to multiple hypothesis corrections for false discovery.

Second, on business performance we find these practice effects translate to improvements in profits and sales. Specifically, we find significant and economically meaningful impacts for firms assigned to *Counseling* and *All Three*, with an increase in profits of 35% (0.27 s.d.) and 21% (0.17 s.d.), respectively compared to *Control*. Likewise, the analysis reveals statistically significant improvements in sales, representing a 17% (0.15 s.d.) improvement for *Counseling*; and a 16% (0.14 s.d.) increase for *All Three*. Businesses assigned to *Movie* also increase their profits and sales with respect to *Control*, but the effects are not statistically significant at conventional levels, partly due to low take-up of the movie. Following the null effects on practices, we find no significant effects on profits and sales of offering the handbook alone. In fact, the three groups with behavioral add-ons significantly outperform the *Handbook* group on both profits and sales.

Next, we turn to mechanisms of impact to understand the pathways for achieving these performance gains. We find no significant increase in the number of customers, total expenses, shop size, or number of employees for any of the treated businesses. Instead, we find improvements in business operational efficiency due to adoption of business practices. Our analysis shows that stocking up and marketing are the two most important channels that mediate the impact on sales and profits. Unpacking the improvements in these practices, we find significant improvements in stocking responsiveness to product profitability, negotiating lower prices with suppliers, consulting with former customers to understand changes in market demand, and offering discounts to loyal customers to maintain customer continuity. Collectively, these practices point to improvements in business efficiency – a mechanism that implies that higher sales can be achieved without an equivalent and corresponding increase in expenses. As a placebo test, we find no significant improvements in business practices that were not covered in the handbook.

Overall, these results suggest both informational and behavioral constraints are jointly binding and that efficiency gains can be achieved by curating and distributing existing and locally relevant knowledge. As hypothesized, we find that information alone – even when curated – is not enough to facilitate behavior change. The movie helps move the needle on performance outcomes we measure, but ultimately the changes are not statistically significant. We also do not detect any significant changes in the growth aspirations of retailers, which suggests that the movie was not effective in this regard. Third, the success of the treatments including counseling suggests that information constraints were present and a nudge in the form of short in-person visits was sufficient to help

retailers internalize the information in the handbook. Inattention, status-quo, and procrastination are likely important channels explaining the success of the counseling nudge, but it is not possible to disentangle among these with the data available in our study. Finally, a simple cost-benefit analysis shows that the highest increase in monthly profits in our study is more than double the per capita cost of the program. Hence, the study points to a highly cost-effective method for achieving efficiency gains.

This paper contributes to at least three literatures. The first is an emerging literature on scalability of programs to facilitate firm upgrading. As McKenzie (2020) highlights in a recent review paper, the biggest challenge for policymakers relates to scaling up good quality business development services to the vast numbers of small businesses that operate in developing countries. Even the best classroom-based business training programs have limited scalability due to cost, infrastructure, and instructor availability constraints. In contrast, our study shows that combining and scaling the handbook with behavioral add-ons can be achieved at relatively low cost, without the need for quality instructors or infrastructure such as classrooms or technical media. As such, our approach offers an alternative and novel avenue for low cost scalability of locally relevant business practices.[2]

Second, our findings are closely related to the literature on firm upgrading and 'PPDs among SSEs' ["Persistent Performance Differences among Seemingly Similar Enterprises"] in a fundamental way (e.g. Gibbons and Henderson (2013)). The motivation of our experimental design is based on the empirical fact that profitable practices exist within the market, yet they do not diffuse without outside intervention. As pointed out by Verhoogen (2020), a key factor preventing diffusion of knowledge is that much of the knowledge needed to perform well is tacit and cannot be simply purchased on the open market.[3] Many organizational capabilities "need to be worked out in the practice of producing," while Gibbons (2010) highlights that they need to be 'homegrown.' (pg. 30). Similarly, Arrow (1969) emphasizes the importance of decoding tacit knowledge to facilitate the diffusion of innovation, arguing that "different communication channels have different costs (or, equivalently, different capacities), where these costs include the ability of the sender to 'code' the information and the recipient to 'de-code' it" (pg. 33). To illustrate the problem of transferring organizational knowledge, Winter (2006) uses the analogy of baking a cake from a recipe: "knowing how to bake a cake is clearly not the same thing as knowing how to bring together in one place all the ingredients for a cake" (cited in Gibbons and Henderson (2013), pg. 38). These are precisely the types of frictions that we address in our study.

Finally, our study is related to the recent literature on mobilizing peer learning to stimulate business growth (Brooks et al., 2018; Cai and Szeidl, 2018; Fafchamps and Quinn, 2018; Lafortune et al., 2018; Hardy Morgan and McCasland, 2016), though there are key differences. In our design, local knowledge is identified, processed, aggregated, and diffused anonymously by us who act as *intermediaries*. By contrast, in Cai and Szeidl (2018), Fafchamps and Quinn (2018) and Brooks et al. (2018) the local knowledge is shared directly and personally by the firm owners who are encouraged to meet up regularly.[4] This not only makes the interventions more costly, but also has

[2]For literature on market based solutions see Maffioli et al. (2020) and Anderson and McKenzie (2020); for online or edutainment delivery see Bjorvatn et al. (2020) and Barsoum et al. (2018); and for firm targeting see Anderson et al. (2018) and McKenzie and Sansone (2019).

[3]The idea of 'tacit knowledge' is not new and dates back to Polanyi (1958), who describes it as the psychological and social construct of overall knowledge.

[4]The information that experienced mentors typically share in Brooks et al. (2018) does not necessarily concern managerial practices of the local context but mostly relate to current information about local market conditions.

the disadvantage that firm owners may choose whom they want to share their best practices with and to what extent. Cai and Szeidl (2018) show that successful business owners share relevant business knowledge with their peers when they are not competitors. Similarly, the findings of Hardy Morgan and McCasland (2016) suggest that competition is an important barrier to the diffusion of technological know-how. In Lafortune et al. (2018) it is former alumni of a formal business training program who share their experiences with peers after having attended the program. In their set-up, the mentors can also choose the extent and type of information they want to share with the current trainees.

Given that competition can hinder the diffusion of knowledge and information as established in these studies, our research design is robust to such competitive effects as it does not involve direct interaction between peers; in fact, we intermediate these transfers of knowledge through the handbook. Moreover, most of the practices covered in the handbook for which we identify positive and significant adoption effects relate to gains in efficiency, and as such are non-rivalrous – i.e. the adoption of a practice by one retailer does not come at the expense of another retailer's performance. This difference may partly explain why we capture larger treatment effects both on the adoption of practices and on business performance. Further, since retailers do not engage in direct contact with their peers, our design allows us to measure the pure effect of information and its different dissemination channels without needing to disentangle the effect of information alone from the effect of contact with peers per-se, which is a usual concern in the peer-to-peer learning literature. Finally, we also differentiate from this literature in that the focus of our study is not manufacturing firms but small retail shops. We believe highlighting the sector is helpful for scalability and external validity as these types of firms are ubiquitous in developing countries and lessons from our study have potentially wide relevance and applicability.

II Research Design

A Hypotheses

Our research is founded on the hypothesis that small businesses face multiple and likely overlapping constraints to the adoption of business practices. This hypothesis offers a potential explanation for the stylized heterogeneity observed in business practices across firms of similar size, which is present even in developed economies such as the US (Bloom et al., 2019). We design and implement a cost-effective and scalable experimental design that identifies the role of informational and behavioral constraints on the adoption of business practices and ultimately on business performance among small-scale firms.

A handbook of business practices forms the basis of our design. The handbook aims to address informational constraints associated with the adoption decision. A key advance in our study is that the handbook identifies and prioritizes particular lines of information that are most relevant for performance improvements in the local setting. In addition, the handbook clarifies common misconceptions about business practices and offers practical tips on adoption. A good analogy for the handbook, following the argument by Winter (2006), is that it provides a curated list of the best local recipes for domestic chefs rather than a comprehensive and universal cookbook. For these reasons, we expect business owners to find it easy to read and follow, which in turn allows us to

focus on low-cost and scalable behavioral nudges.

The behavioral nudges of the study are derived from the hypothesis that information alone, even if curated, may not be sufficient to facilitate sustained behavior change. The literature identifies complementary constraints related to inattention (Steiner et al., 2017), aspirations (Dalton et al., 2016), status-quo bias (Kahneman et al., 1991), and procrastination (Duflo et al., 2011) that can individually or collectively block behavior change among informed decision-makers. With this motivation in mind, our study design complements the handbook with two behavioral add-ons that seek to relax these additional constraints.

The movie targets the aspirations constraint, and is based on the notion that the handbook may not be helpful if business owners do not aspire to grow. The format of the movie used in our study with successful peers implementing practices is motivated by recent evidence that such media can influence aspirations and behavior (Bernard et al., 2014; La Ferrara, 2019). Acting as role models, these peers, along with their unique success stories and hands-on advice, may serve as an exemplar and therefore foster growth by opening up aspiration windows and facilitating the adoption of successful practices.

The in-person counseling targets potential constraints related to inattention, status-quo bias, and procrastination over information alone. Moreover, the psychology literature has long advocated the benefits of human interaction in counseling over inanimate information sources such as computer messages or printed materials (King et al., 2007). Based on this evidence, we consider the short facilitator visits as a nudge that can help internalize the lessons of the handbook.

Finally, the combined treatment of both the movie and counseling coupled with the handbook (*All Three*) is designed to alleviate the information as well as both types of behavioral constraints discussed above. Moreover, this study arm seeks to identify possible complementarities across treatment arms.

We acknowledge that both the movie and counseling can influence behavior through different and possibly complementary channels than those hypothesised above. The movie may serve as an endorsement of the handbook by successful peers, or it may further facilitate reading of the handbook as it constitutes an accessible means of communicating key ideas. Likewise, the counseling intervention may facilitate learning on top (or instead) of acting as a nudge, since it offers a) additional tailoring to individuals' specific circumstances, b) one-to-one interaction on topic areas that are elicited prior to each session, and c) trained facilitators who themselves come from similar backgrounds as the retailers. Although we are not able to separate these channels experimentally, in Section B we analyze available survey evidence to try and discern whether any of these channels are activated in our study sample.

B Study Setting

This study is based in Jakarta, Indonesia. With a population of 10.1 million in inner Jakarta and an urban area of around 30 million ("DKI Jakarta"), Jakarta is the largest city in South-East Asia, and the capital and economic center of Indonesia. In 2015, the city generated a nominal GDP of almost one-sixth of the total nominal GDP of Indonesia (Statistics Indonesia, 2016).

Our sample consists of traditional retail shops, locally called *Warung* or *Toko Kelontong*. Most of the shop owners in our sample are situated in residential areas or adjacent to "wet markets" for

meat, fish, and vegetables.[5] Our baseline analysis shows a spread of 24 different product categories offered by shop owners and the average shop owner selling a variety of them. 71% of the sample lists tobacco and cigarettes among their top-3 most sold products and 50% list this category as their top selling product. Out of the shops that do not sell tobacco and cigarettes; rice, gas and petrol, and soft drinks are the main products on sale. Less than 5% of shops list fruits and vegetables or meat and fish as their top-3 main selling products. 9% of the sample list eggs in their top-3, but only 2.5% list it as their top product.

Small retail businesses make up a large fraction of all micro and small enterprises (MSEs) in Indonesia: about 22% of all employees in MSEs work in retail and hospitality which makes it the second largest sector after agriculture (Indonesian Ministry of Cooperatives and SMEs Indonesia, 2011). The sectoral choice of our study is helpful for scalability and external validity as these types of firms are ubiquitous in developing countries.

C Sampling

The city of Jakarta comprises 144 districts ('Kelurahan'), which include the urban area of Jakarta proper as well as some agglomerations in the wider Jabodetabek metropolitan area. We restricted the sampling base for this study to the 112 districts in South, East and West urban Jakarta.[6] As a first step, 29 of the 112 districts were randomly selected to be included in the study.[7] Across the 29 selected districts, we first conducted a listing exercise of small retail shop owners that met the following three inclusion criteria: (1) the shop is at least $4m^2$ in size; (2) the shop offers at least two different product categories; and (3) the shop is at least 30 meters away from other shops in the listing.[8] In addition, we excluded movable establishments, franchise businesses of larger retail chains, and shops located in densely populated marketplaces.

These listing criteria were chosen to ensure the sample would consist of business owners with an established store, who sold a variety of products, and where spillovers were minimized by design. Out of the 2,042 businesses listed through this method, we randomly selected 1,301 to be part of this study.

D Experimental Design and Timeline

The 1,301 shops in the sample were randomly divided into a handbook treatment group ($N = 1040$) and a pure control group ($N = 261$), stratified by district, gender, shop size (below $6m^2$, between 6 and $10m^2$, or above $10m^2$), and a composite score of business practices (above and below the median). Among handbook recipients, we implemented two additional and orthogonal treatments: the movie screening and the two counseling visits. Appendix Figure 2(a) maps the spatial distribution of our study sample across Jakarta and Appendix Figure 2(b) illustrates the spatial distribution in one sample district.

[5] Appendix Figure 1 shows pictures of two typical shops in our sample, representative in both size and appearance.

[6] We excluded all 32 villages of North Jakarta due to a MSE training program being run concurrently in that area by a local retail chain.

[7] We initially selected 30 districts, however, in one of these districts only five businesses were identified and they differed markedly from the remaining sample. Hence, they were dropped from the sample.

[8] Within a market, we picked a starting point (i.e. the first shop in the sample) at random. From here, we made sure that all subsequent shops in the sample were at least 30 meters away from the first shop and from each other. This protocol ensured that we have a minimum of 30 meters distance between shops throughout the study sample.

Overall, the study design consists of four experimental treatment arms of 260 firms each:[9] handbook only (*Handbook* group), handbook and an invitation to the movie screening (*Movie* group), handbook and an invitation to two counseling visits (*Counseling* group), and all three interventions (*All Three* group).[10]

The timing was as follows.[11] In January 2016 we sampled the 2042 businesses. In March and April 2016 we administered the baseline survey and registered the trial of the study at the American Economic Association's Randomized-Controlled-Trial Registry website.[12] Interventions took place in October and November 2016 and were followed by a midline survey held in April and May 2017 and then an endline survey held in May 2018.

D.1 Handbook

The content of the handbook was developed through primary qualitative and quantitative assessment of best practices among local business peers. First, prior to the start of quantitative fieldwork, we conducted detailed qualitative interviews with a sample of 102 shop owners chosen from two comparable urban markets outside our study area. The goal of these interviews was to understand the most common and successful business practices from the perspective of local business owners, and to familiarize ourselves with the various implementation processes and constraints to adoption. With this objective in mind, we asked a number of open-ended questions in the format of a conversation, and responses were recorded and later transcribed.[13]

The analysis of these interviews allowed us to identify the following five categories of locally relevant business practices: record keeping, financial planning, stocking up, marketing, and joint decision making. In the next stage, these categories directly fed into the quantitative baseline survey, from which we were able to associate the contribution of within-category individual practices to profits, sales, and number of customers. Specifically, for each category we estimated linear regressions of profits, sales, and number of customers on individual practices and a set of firm-level controls. Appendix Table I2 shows the (top) practices and the baseline regression coefficients. Column 1 presents the number of specifications in which the variables had a significant coefficient and Columns 2 and 3 present the coefficients of the regressions with sales and profits as dependent variables, respectively. These regression coefficients were used in the handbook to illustrate a quantitative association of a particular practice with sales and profits. The findings from these quantitative assessments were then matched with the qualitative fieldwork to arrive at the final set of locally relevant best practices to be included in the handbook. With this information and the help of a local NGO, we developed a handbook of best practices implemented by peers. The handbook consists of five chapters: keeping business records, calculating profits, making stock-up decisions, attracting customers, and cooperation in business decisions, in this order. Although there is no chapter exclusively on financial planning, practices such as reviewing financial performance, analyzing where there are areas for

[9] Appendix Figure 3 tabulates the various treatment arms of this study

[10] Among the 1,040 handbooks, 520 had the economic returns to the adoption of each business practice described as gains and 520 had them described as losses. However, as shown in Appendix G and Appendix Table G1, the effects of the framing on the main outcomes of this paper are statistically indistinguishable except for the *Movie* group. For this reason and to maximize statistical power, we focus on pooled estimates in this paper.

[11] See Appendix Figure 4 for a detailed timeline.

[12] https://www.socialscienceregistry.org/trials/1175. Appendix L provides a report clarifying any departure of the research analysis in this paper from what was pre-registered in the AEA RCT Registry.

[13] Respondents were informed that the interviews would be used to develop a program to help small businesses and to advance research in Indonesia.

improvement, and comparing sales with targets are natural consequences of implementing the best practices suggested in the handbook. Appendix Table I1 lists all the practices measured in this formative stage of the study, and Appendix K presents a summary of the information on beliefs, reasons to adopt practices, step-by-step implementation guidance, and tips presented in each of the five chapters of the handbook.[14]

Each chapter of the handbook is structured as follows. First, it confronts common false beliefs and misconceptions held by shop owners about the usefulness of implementing the different practices. These misconceptions are based on statements that were heard repeatedly during the qualitative interviews. For example, one common thought by shop owners was that keeping records is difficult for individuals without higher education, or that it is complicated. The handbook emphasizes that this is not the case, and that we have observed shop owners from different educational backgrounds keeping records. Likewise, to confront the belief that keeping records is complicated, the handbook provides simple step-by-step guidance on how to make record keeping an easy routine to follow, also learned from peers.

Second, it presents arguments for why it is important to implement the practices, providing locally relevant evidence-based reasons. To this end, the handbook uses regression coefficients associating top practices with profits and sales using data from the quantitative baseline survey. This information is presented in simple layman terms.[15] As a concrete example, consider the marketing practice of consulting former customers, which is a binary variable equal to 1 if the retailer consulted with former customers on why they stopped buying from this shop. As Appendix Table I2 shows, the baseline analysis returns coefficients of 0.24 on sales and 0.23 on profits. On page 7 of handbook, this information is conveyed as follows: "retailers who decide to get back to former customers to see the reasons why they quit buying at their shop have monthly sales 24% higher than the sales of businesses whose owners just let it go. Their monthly profits are also 23% higher."

In addition, the handbook provides an approximation to the monetary value of implementing each practice, both for a "typical" size shop and for "larger" shops. The following is an illustrative example: "From the survey we know that shops that use discounts to attract new and retain loyal customers have monthly sales that are 40% higher than the sales of shops which do not give discounts. Also, their monthly profits are 29% higher." And this is the monetary value: "For a typical shop with IDR 15 million in monthly sales, shops that offered discounts earned IDR 21 million in sales. For a bigger shop with sales of IDR 30 million per month, it would mean IDR 42 million in monthly sales." Finally, the handbook suggests other reasons to implement the practices. For instance, it suggests that profit-calculation "is useful to plan finances and save up money" and record keeping is necessary to "compare your sales with a benchmark you have", or to "be better able to save up and withstand unexpected events', or "unless you keep proper records, there is no good way of knowing how much cash you have on your hands or how much to save. And without savings, unexpected events can hit you and your family hard."

Third, the handbook provides clear, step-by-step implementation guidance illustrated by idiosyncratic practical examples taken from exemplary shop owners. For example, record keeping involves nine specific steps, stocking up comprises seven steps, and marketing involves three steps. The

[14]The full handbook can be found here.

[15]To make sure retailers understand these messages, we developed the handbook after several pilots with retailers of similar backgrounds in Jakarta, and they found this type of information comprehensible. Moreover, the average retailer in our sample has 9 years of formal education, which implies that they possess basic reading and mathematical comprehension.

following is an example that we used in the handbook to introduce the practice of joint decision making: "Imagine you are in the back of the shop making a list of items to stock up while your spouse is serving customers in front. You should ask your spouse which items are selling well. That way, you will know better what exactly to stock up on and by how much."

Fourth, the handbook offers several tips to facilitate the adoption of practices that we learned during the qualitative interviews with shop owners. For example, it explains how to deal with the uncertainty of electricity bills in Jakarta, and why switching to a local voucher system works well for a lot of firms. Regarding marketing practices, one tip is to take advantage of the fact that suppliers may want to leave a new product in the shop without charging for it right away. For instance, in the handbook we state: "Remember that the suppliers also want to find out how well the new product sells in the market." As another example, in the handbook we refer to the following tip to induce joint decision making: "by including your co-worker or spouse in the process of making decisions about the shop, chances are they will become more involved and thus be of even greater use." For record keeping, the handbook recommends using different colors for household expenses, family "loans", monthly expenses (e.g. bills, stock up, etc.), and brackets for stalled payments. For stocking up, the handbook suggests using tallies.

To finalize the description of the handbook, it is important to highlight that curating local knowledge in this manner is quite scalable. We worked with an initial sample of 102 firms (less than 10% of our study sample), and gathered rich qualitative data from them. The variance within this sample was very informative for the purposes of the handbook.[16]

D.2 Documentary Movie

The movie aims to inspire and motivate the adoption of business practices with the help of locally selected successful business peers. The motivation for this intervention is grounded in research that suggests that exposure to positive role models can facilitate behavior change (Chong and La Ferrara, 2009; La Ferrara et al., 2012; Berg and Zia, 2017; Riley, 2017) and ignite aspirations and forward-looking behavior (Bernard et al., 2014).

The retailers who participated in the movie were selected from the initial pool of 102 qualitative survey participants. Based on these interviews, we shortlisted nine retailers who were ostensibly successful, had an organized shop and books, and employed the largest number of business practices in our study. These characteristics qualified the selected business owners as potential role models who could be inspirational for their peers and alleviate aspirational constraints.

In conjunction with the production of the handbook, we conducted further in-depth interviews with the selected owners about their personal business trajectory and about business practices and implementation advice they regarded as crucial to achieve growth. We also wanted diversity in gender, age, and ethnicity to appeal to the different business owners in our sample. This heterogeneity is important since similarity cues based on these factors have been shown to facilitate social learning besides cues of success, competence, skill, and knowledge (see, e.g., Rendell et al., 2011; Efferson et al., 2008; Chudek et al., 2013; Henrich and Gil-White, 2001; Corriveau and Harris, 2009; McElreath et al., 2008).

This procedure resulted in a final set of five shop owners representing the local frontier of best

[16] In Appendix A we provide a detailed comparison of the handbook vis-a-vis different approaches to business training found in the literature.

practices in each domain of business operations. Moreover, these shop owners regularly employed the practices in the same way featured in the handbook and agreed to describe their implementation methods and paths to success in a documentary movie. The movie featuring these five successful business owners was filmed on shop sites and edited in post-production by a professional media company hired by us. We were involved at each stage of implementation, including script development, test runs, filming, and post production. The end product was a 25 minute movie.[17]

The movie was publicly screened in each of the 29 districts at a local school or other public space. All screening locations were central and accessible to all invited businesses. In order to increase attendance, shop owners were offered IDR 100,000 (USD 24.68 PPP) as a show-up fee and to compensate them for their time and transport expenses. In addition, we offered two alternative screening dates in each district and sent individual text message reminders the day prior to each screening. Each screening was followed by a facilitation session by a trained counselor who clarified any doubts and answered questions from the audience. The screening ended with a short feedback survey and payment of the show-up fee.

D.3 Counseling

For the counseling treatment, we trained a set of local facilitators based on the content of the handbook. These individuals were required to have a bachelor's degree in a related field and some experience interacting with businesses similar to those in our study. All facilitators were fieldwork enumerators who expressed interest, had the experience of conducting business interviews, but were not involved as enumerators in the survey-stage of this study. The training of the enumerators was conducted over three days and included classroom lectures, role play exercises, and pilot visits to retail businesses in districts external to the study. The 20 facilitators trained through this process were then randomly assigned to businesses in our study and were supervised by senior field staff.

The objective of the counseling treatment was to reduce procrastination and inattention associated with reading and applying the content of the handbook. We aimed to achieve this objective by implementing the following protocol. The facilitator first confirmed the identity of the business owner and then asked which aspects of the handbook needed clarification. Based on the owner's response, the facilitator chose one of three options. First, if the entrepreneur had started implementing a practice but encountered problems along the way, the facilitator would document the issues and start giving standardized implementation advice. Second, if the entrepreneur had not started implementing any practice but had made progress reading the handbook, the facilitator would document any issues with the material and then give implementation advice. Once all issues were dealt with, the facilitator would encourage the entrepreneur to go through the rest of the chapter and follow the written guidance. Third, if the entrepreneur had not even started reading the handbook, the facilitator would elicit their priorities among the practices and start introducing the chapter corresponding to the most relevant practice.

Each shop visit lasted approximately 30 minutes. At the end of the first visit, the shop owners were asked to establish goals for the implementation of a practice covered during the visit and for the study of selected material. A second visit was scheduled two weeks after the first and at the convenience of the shop owners. This second visit followed the same protocol as the first with the

[17] The movie can be accessed here.

difference that the starting point was determined by the work left from the first session and the shop owner's priorities elicited during that visit.

There are several features that make the counseling sessions special. First, the sessions were provided on-site rather than in an outside venue, which allowed shop owners to tend to their business matters with only minor interruption. Second, the counseling sessions involved direct one-to-one interaction on topic areas that were elicited prior to each session. As a result, although the sessions were only 30 minutes each, they were intense in engagement. Third, the facilitators themselves came from similar socioeconomic backgrounds as the shop owners, which was a deliberate choice to facilitate closer interaction.

III Data

Firm-level data were collected at baseline, mid-line (6 months after the intervention), and end-line (18 months after the intervention). The baseline assessed business owner background characteristics, business characteristics, and the use of business practices.[18]

A Summary Statistics, Attrition, and Compliance

Summary statistics from the baseline survey are presented in Table 1. Columns (1)-(2) provide mean and standard deviation values for the total sample of 1301 businesses, respectively, while column (3) presents p-values for equality of means tests across all experimental groups (treatment and control). All p-values are large (well above 0.10). We also ran a Multinomial Logit specification regressing assignment to treatment on baseline characteristics and checked for joint orthogonality with a Chi-Square test. The p-value of this test is 0.819, which also suggests that the randomization was successful.

According to the statistics in Table 1, shop owners in our sample are mostly female (71%) and are 45 years old on average. Educational backgrounds are mixed, with a mean educational attainment of 9 years of schooling. The average entrepreneur is risk averse, and is neither patient nor impatient.[19]

The average business has a size of about 13 sq. mts, has been in operation for 13 years, employs two workers, receives about 50 customers per day and has monthly sales of 4786 USD PPP and monthly profits of 908 USD PPP (rounded).[20] Only 19% of the businesses report having a tax ID and 16% obtained a loan in the last 12 months. The average adoption rate of business practices included in the handbook is 36% at baseline, compared to 18% for practices not covered in the handbook. Finally, the normalized index for shop appearance as judged by the enumerators is 0.63, where 0 indicates the least tidy and 1 the most. This empirical observation implies that the shops were considered to be in relatively good condition.

[18] Appendix Table I1 shows the list of the business practices measured in the surveys.

[19] Risk attitude is measured with the answer to the question: "Some people usually avoid taking any risk, others are generally fully prepared to take risks. Please imagine a yard stick from 0 to 10. This time 0 means you usually avoid taking any risk and 10 means you are generally fully prepared to take risks." Time preferences are measured in a similar way: "Some people usually want to have things now rather than later, others are generally willing to wait a long time. Now, please imagine a yardstick from 0 to 10. 0 means you usually want things now rather than later and 10 means you are generally willing to wait."

[20] The profit and sales figures are winsorized on both tails, at the 5^{th} and 95^{th} percentile. Profits are calculated by aggregating all costs up and then calculating sales minus total costs. This applies some of the suggestions on how to calculate more accurate profits from Anderson et al. (2019). Appendix E explains the estimated profits measure used in this paper, and for robustness presents regression results using alternative measures.

This sample of retailers is similar on many observable characteristics to other samples compiled by McKenzie and Woodruff (2017) in Bangladesh, Kenya, Mexico, Ghana, Nigeria, Sri Lanka, and Chile. Moreover, MSEs like the ones in our sample constitute the majority of small firms in the developing world. According to Hsieh and Olken (2014), "about 90 percent of firms in Mexico employ less than 10 workers. In India and Indonesia, the fraction of firms with less than 10 workers is almost visually indistinguishable from 100 percent" (p.93). Similarly, McKenzie (2017) reports 99.6 percent of firms in Nigeria have fewer than 10 workers.

Online Appendix B presents the analysis on survey attrition. As Appendix Table B1 shows, attrition was very low. We were able to reach 92% of the sample at midline and 81% of the sample at endline, and the small attrition rates are not correlated with treatment status. Nevertheless, Appendix Table B2 presents Lee bounds for the main treatment effects in this paper and finds statistically significant impacts even at the lower bound.

Online Appendix C discusses treatment compliance and attendee evaluation, summarizes statistics collected during the counseling visits, and presents analysis for selection into treatment. Appendix Table C1 shows that out of the 520 shop owners invited to the movie screening, only 52% showed up at the venue for the film screening session. The take-up of the counseling treatment was higher - out of the 520 shop owners offered personalized counseling sessions, 77% received the counseling once and 68% received it twice. Lastly, we do not observe consistent discernable patterns of selection on observables (Appendix Table C3).

IV Main Results

A Estimation Strategy

We present pooled intent-to-treat (ITT) estimates using the following ANCOVA regression specification:

$$Y_{(2,3)i} = \alpha + \sum_{m=1}^{4} \beta_m \mathrm{T}_{mi} + \gamma \mathrm{X}_{1i} + \delta V + \theta M + \zeta \mathrm{Y}_{1i} + \epsilon_i, \tag{1}$$

where $Y_{(2,3)i}$ is the stacked outcome for business i at midline $t = 2$ and endline $t = 3$.[21] T is a dummy variable equal to one if business i was assigned to a particular treatment group, while $m = 1$ to 4 represent the four types of interventions in the study. Since the randomization was done after stratifying by gender, shop size (micro, small, or mid-sized), and a median split of a business practice composite score, we follow Bruhn and McKenzie (2009) and include the strata dummies represented by the vector X. V represents district fixed effects, while M is a dummy variable for the midline survey round. Y_{1i} is the baseline value of the outcome of interest. Standard errors are clustered at the business level.[22]

[21] Appendix Table H6 reports regression results on practice adoption, profits, and sales after 6 and 18 months using the main ANCOVA specification in eq. 1. Given that we do not see differential effects across waves, we present pooled estimates.

[22] We did consider Treatment-on-Treated (TOT) analysis owing to the low take-up rates especially for the movie. The TOT coefficients do scale up by the inverse of compliance, but standard errors also increase so the p-values are not different or significant for TOT.

To correct for multiple hypothesis testing we adopt the method of Benjamini et al. (2006) as outlined in Anderson (2008) to calculate sharpened FDR q-values.[23] This method applies the (Benjamini and Hochberg, 1995) (henceforth BH) correction in two stages where the first stage uses the standard BH procedure to estimate the number of true null hypotheses and the second stage uses these estimates and re-applies the BH procedure to achieve sharpened FDR q-values. This procedure provides better power than the standard BH method (Anderson, 2008). Given the multiple treatment arms and multiple outcomes in our study, we balance the issue of power against the likelihood of false discovery by adopting the following approach. First, we group outcomes into domains, where each domain encompasses a particular aspect of the business. Second, we apply FDR corrections simultaneously for all tests performed within each domain.

We specify five such domains: business practices, business performance, business expenditure, mechanisms, and personal attributes of the business owner. The business practices domain includes aggregate indices for record keeping, planning, stocking up, marketing, and joint decision making. We also present an overall aggregate index of practices and compare it against alternative aggregation techniques such as principal component analysis and lasso in Appendix D. The business performance domain includes profits and sales, which are the main performance outcomes of a firm. We also construct an overall performance z-score measure, where we first compute z-scores for profits and sales and then take the average of the two z-scores. The expenditure domain includes total business expenses as well as disaggregated individual expense categories. The mechanism domain includes shop size, total number of employees, and total number of customers, i.e. outcomes that inform on the mechanism of treatment impact. As with the performance domain, we construct an overall mechanism z-score measure. Finally, the personal domain includes business owner outcomes of satisfaction with life, satisfaction with finances, systematic and intuitive working style, as well as an index of standardized z-scores for business aspirations.[24]

B Business Practices

We first analyze primary treatment impacts on business practices that were covered in the handbook. In Table 2 we report results for the five dimensions of business practices and the aggregate of all practices with sharpened FDR q-values reported in italics.[25] The table shows that providing the handbook alone leads to no significant improvement in the adoption of business practices in any category. The point estimates are also very small and close to zero. In contrast, we find strong and statistically significant impacts when this pure information treatment is combined with the behavioral add-ons.

First, we find that adding either the movie, counseling, or both results in statistically significant improvements in the aggregate of all practices. Compared to an adoption rate of 33.7% in *Control*, our results show a 12.5% improvement (0.22 s.d.) for *Movie*, a 10.1% improvement (0.18 s.d.) for *Counseling*, and a 16.0% improvement (0.29 s.d.) for *All Three*. All three coefficients are statistically significant at the 1% level. Furthermore, these coefficients are statistically larger than providing the *Handbook* alone, as represented by significant F-test p-values at the bottom of column (1).

[23] The FDR is the expected proportion of rejections that are type 1 errors (i.e. false rejections).

[24] The aspirations index is an average of z-scores for 12-month respondent aspirations for shop size, employment, customers, and sales. Given that the satisfaction questions are only asked in the endline and the cognitive style questions are only asked in the baseline and midline, this domain does not have an overall z-score measure.

[25] Online Appendix D presents regressions results for individual business practices in Appendix Tables D2-D6.

Second, disaggregated analysis in columns (2)-(6) shows that even with FDR corrections for all tests performed across practice categories and treatment arms, we find that the *All Three* intervention significantly stimulates the adoption of business practices on all dimensions of the handbook. Specifically, the results show statistically significant impacts on record keeping (27.6%), planning (16.9%), stocking up (11.3%), marketing (24.4%), and joint decision making (21.9%).

We also analyze the distributional impact on business practices by plotting and comparing the kernel densities of the business practices aggregate for each treatment group. Figure 1 plots the Probability Density Functions (PDFs) for the baseline and endline surveys. We then perform Kolmogorov-Smirnov tests of equality of distributions for all treatment and control arms at baseline and endline. For baseline, we cannot reject equality of distributions for any combination, which is by design as the assignment to treatments was random. For endline, we can reject the equality of distribution null between the *All Three* treatment group and the *Control* at the 5% significance level. Appendix Table D7 presents quantile regression analysis for each decile of the business practice distribution and finds statistically significant effects in the *All Three* treatment group at every decile except the highest. These results show that the treatment was effective in improving practices among all businesses, even those at the lower tail of the business practices distribution.

Finally, as a placebo test, we measure impacts on unrelated business practices. Specifically Appendix Table H2 repeats the analysis for practices that were not covered in the handbook, and finds no treatment impacts on either the composite or any of the individual practices. This finding implies that the practices that we intended to treat indeed exhibited a change in the treatment groups while practices unrelated to our intervention were not affected.

C Business Performance: Profits and Sales

Next we turn to performance impacts of the handbook and behavioral add-ons in Table 3. Column (1) presents results for the average z-score of profits and sales; column (2) for the estimated monthly profits, and column (3) for the monthly sales. Both profits and sales are winsorized on both tails at the 5^{th} and 95^{th} percentiles. Appendix E explains how the profits measure is computed and shows that the reported impacts are robust to alternative profit measures.

Table 3 shows statistically and economically significant treatment effects on the three performance measures in the *Counseling* and *All Three* groups. Compared to *Control*, businesses assigned to *Counseling* improve profits by 35%. This effect is statistically significant at the 1% level and represents a 0.27 standard deviation increase with respect to *Control*. Businesses in *Counseling* also increase their monthly sales by 17% (0.15 standard deviations) and the effect is significant at the 5% level. Similarly, businesses assigned to *All Three* improve profits by 21% and sales by 16% (0.17 and 0.14 standard deviation improvement) compared to *Control*, effects that are statistically significant at the 5% level.

In monetary terms, these effects are sizable, and in the *Counseling* group translate to USD PPP 310 more profits on a monthly basis and a USD PPP 835 increase in monthly sales. In *All Three* the outcome effects imply USD PPP 190 increase in monthly profits and USD PPP 803 increase in monthly sales. Table 3 also illustrates that businesses assigned to the *Movie* improve profits by 13% (0.10 standard deviations) over the *Control* group, but this effect is not statistically significant at conventional levels. The lack of statistical significance is in part due to the relatively low take-up of the movie.

The impact of *Counseling* and *All Three* on profits is quite large, especially considering that firms only received one hour of counseling in total. To put this effect size into context, other studies that show positive impact on profits find comparable size effects, but utilize extensive "treatment hours" (see Appendix Table A14 for a comparison with other studies). Specifically, with respect to the profit impact, de Mel et al. (2012) report an increase by 43% after 49 to 63 hours of training; Calderón et al. (2013) find a 24% increase after 48 hours of training; Lafortune et al. (2018) find a 31% increase following 48-56 hours of role model training; and Cai and Szeidl (2018) find a 35% increase after 144 hours of owners-managers meetings.

Table 3 also shows that the handbook alone does not have any significant impact on profits or sales; in fact, the F-test p-values in the bottom half of the table show that the *Movie*, *Counseling*, and *All Three* significantly outperform the *Handbook* on all performance outcomes. These findings suggest that the curated information provided in the handbook is internalized only after it is accompanied with behavioral add-ons in the form of counseling or both counseling and the movie.

Finally, we analyze the distributional impact on profits and sales by plotting the Probability Density Functions (PDFs) of estimated profits and sales at baseline and endline (see Figure 2). We also perform Kolmogorov-Smirnov tests of equality of distributions for all treatment and control arms at baseline and endline. While we cannot reject equality of distributions at baseline, we reject the equality of distribution null at endline between the *All Three* treatment group and *Control* at the 5% significance level for estimated profits and 1% level for sales. In addition, for estimated profits at endline, equality of distributions is rejected between: (a) *All Three* and *Handbook* at the 5% level; (b) *All Three* and *Movie* at the 10% level; (c) and *Counseling* and *Handbook* at the 10% level. For sales at endline, equality of distributions is rejected between: (a) *All Three* and *Handbook* at the 5% level; (b) *Counseling* and the *Control* at the 1% level; (c) and *Counseling* and *Handbook* at the 10% level.

Finally, to benchmark the findings of our study, we compare the standard deviation effect sizes on profits and sales with what other studies in the literature have found. As Appendix Table J1 shows, most studies in this literature do not report control group standard deviations, so this comparison has limited scope. There is also variation in how profits and sales are estimated in the literature as measuring business performance still remains a challenge. Nevertheless, the two studies for which these data are available are Anderson et al. (2018) and Lafortune et al. (2018). Following 80 hours of intensive business training on a pre-screened sample of aspiring entrepreneurs, the former study finds a 0.3 SD increase in profits and a 0.3 SD increase in sales. Following 48-56 hours of role model based training and individual technical assistance, the latter study finds up to a 0.17 SD improvement in profits and up to a 0.19 SD improvement in sales. In comparison, we find a 0.15 SD improvement in profits and 0.14 SD improvement in sales for *All Three*. The handbook involved no beneficiary interaction, the movie was 25 minutes long, and the two counseling sessions were 30 minutes each. Notwithstanding the substantial difference in hours of involvement with beneficiaries, the effect sizes we find are still consistent and comparable with what has been identified in the literature.

D Business Expenses

Table 4 presents regressions results for total expenses in the last month in column (1) and for individual expenses (columns 2 to 6). While the regression analysis does not show statistically significant impacts, the results in Table 4 do show higher point estimates for precisely the treatment

groups that have corresponding and significant increases in sales. The coefficients are positive and directionally consistent, but the standard errors are large and therefore the coefficients are not statistically significant at conventional levels. Nevertheless, column (1) of Table 4 shows that total expenses in *All Three* are significantly larger than *Handbook*, with a p-value of 0.09 in an equal coefficients F-test. Decomposing this effect into individual components, Table 4 also shows that stocking up expenses account for more than 95% of total business expenses when comparing the control group means across expenditure categories.

E Business Size, Employees, and Customers

Table 5 presents regression analysis for the mechanism domain, which relates to shop size (column 2), number of employees (column 3), and number of customers (column 4), as well as the average z-score of these three outcomes (column 1). Although the coefficients for corresponding treatments with behavioral add-ons are positive, they are not statistically significant for any outcome. These results imply that the increase in sales is not due to new customers or business expansion. In Section V, we explore the mechanisms of impact and tag these improvements as efficiency gains.

F Business Owner Personal Attributes

Before turning to mechanisms of impact, Table 6 presents results on business owners' personal attributes. In particular, we analyze effects on satisfaction with life (column 1) and finances (column 2), on a measure of systemic (column 3) and intuitive (column 4) working style, and on an aspirations z-score index (column 5).[26]

Our hypotheses regarding the impact of the interventions on these personal attributes were as follows. Regarding satisfaction with life, we left it as an open empirical question, since it is not clear that improvements in sales and profits would necessarily imply greater satisfaction, as they may have come along with other subjective costs. On cognitive styles, our hypothesis was that the entrepreneur would think more systematically than intuitively after improving business practices. Finally, because the treatments (especially the *Movie*) had an aspirational component, our hypothesis was that treated retailers would raise their aspirations. However, as Table 6 shows, we do not detect significant treatment effects on any personal outcome, which suggests that the channel practice adoption and performance improvement is purely through gains in knowledge rather than aspirations, working style, or subjective well-being.

[26] We use the following two questions from the World Values Survey to measure subjective well-being: On a scale from 1 ("very dissatisfied") to 10 ("very satisfied") (1) "How satisfied are you with the financial situation of your household?" and (2) "How satisfied are you with your life at this point?" (see, Inglehart et al., 2014). To measure cognitive style, we use a 10-item questionnaire proposed by Sagiv et al. (2010). This includes five statements measuring a systematic approach (e.g., "Before I do something important, I plan carefully.") and five statements measuring an intuitive approach (e.g., "I often follow my instincts.") to working and thinking. The aspirations index is the average z-score of 12-month aspirations for shop size, employment, customers, and sales. These aspirations measures are closely related to those used in the literature (e.g. (Bernard et al., 2014)), although we acknowledge that aspirations are difficult to measure precisely. Columns (1) and (2) outcomes are only available for the endline; and columns (3) and (4) outcomes are only available for the baseline and midline.

V Mechanisms

We now turn to the mechanisms behind the increase in profits and sales identified in the previous section. The experiment design of this paper implies a strong causal link between improved business practices and better performance outcomes. This link is especially merited given the handbook and its accompanying behavioral add-ons are precisely based on the business practices for which we identify treatment effects. In this section, we delve more deeply into the details of these practices and how they relate to gains in efficiency rather than other mechanisms. We also investigate possible alternative explanations for the performance improvements.

A Measurement Concerns

A.1 Efficiency Gains or Just Better Record-Keeping?

In this subsection, we unpack the individual practices underlying the aggregate indices to shed light on the mechanisms of impact. This exercise also helps address the concern that the performance gains we identify are simply the result of shop owners being able to record sales and profits better without any real improvements in the business.

First, it is important to note that there is no effect on profits for the *Movie* treatment (Table 3) even though record keeping improves in that treatment arm. Second, while record keeping practices do improve as a result of the interventions, shop owners improve other key handbook practices as well. In particular, they adopt practices that are bound to make the business more efficient. For instance, businesses assigned to *All Three* "adjust stocks based on product profitability" and "negotiate lower prices with suppliers" (Appendix Table D4); "consult with former customers" and "offer discounts" (Appendix Table D5); "make joint decisions" and "draft agreements to make joint decisions" with a business partner (Appendix Table D6). These practices are expected to either lower the average unit cost of the shop owner (vis-a-vis suppliers) or lower frictions when reaching customers and improve business efficiency. Additionally, we observe a substantial positive impact on practices related to financial planning (Appendix Table D3), such as "reviewing financial performance to identify channels of improvement", "making anticipated budget for upcoming costs" and "comparing target vs actual sales". Combined, these practices strongly suggest gains in efficiency as a primary mechanism for improvement in profits and sales. This efficiency mechanism implies that improvements in sales can be accommodated without an equivalent increase in expenses.

The improvements in individual practices can be traced directly to the content of the handbook, where, for example, the chapter on stocking up does not focus on "stocking up more", rather it focuses on "stocking up efficiently" based on set schedules and restocking targets. Specifically, page 40 of the handbook encourages business owners to stock up based on the demand of customers on the best sold products. Also, pages 41-46 aim to improve the business owners' awareness and management of shop inventory. In particular, the handbook provides step-by-step guidance on how to keep track of inventory and avoid supply chain inefficiencies.

Second, we apply formal causal mediation analysis to investigate the proportion of the treatment effects on sales, profits, and expenses in *All Three* that can be attributed to each dimension of practices. Specifically, we follow the decomposition framework outlined in Carpena and Zia (2020) and motivated by Imai et al. (2011), which separates the average treatment effect on sales, profits, and expenses into an Average Causal Mediation Effect (ACME) and an Average Direct Effect (ADE).

The ACME isolates the impact of a particular intermediary channel (e.g. marketing practices), while the ADE represents all other pathways. Empirically, these effects are estimated using coefficients from two regressions: one for the effect of the treatment on the mediator (i.e. practice score); and the other for the effect of the mediator on the outcome conditional on the treatment. The product of these two coefficients, the ACME, captures the portion of the average treatment effect that can be attributed to the mediating practice.[27]

Appendix Table F1 presents the ACME analysis for profits (panel A), sales (panel B), and expenses (panel C) in *All Three*. With a focus on relative rather than absolute coefficients, the results suggest that record keeping is in fact not the most important mediator of performance, rather stocking up and marketing practices contribute the most to sales and expenses and where the mediation effects are largest and statistically significant. Likewise, for profits, the mediation effects are strongest and statistically significant for marketing practices.

A.2 Real Adoption of Practices or Self-Reporting Bias?

A possible concern with the analysis in this paper is related to social desirability bias or experimenter demand effects. Specifically, the concern relates to treated shop owners simply misreporting higher adoption rates, profits, and sales. First, we would like to note that this is not the first paper facing such a concern; in fact, previous research has specifically tested for self-reporting bias in field experiments and does not find much evidence. For example, Bruhn et al. (2018) compare survey data with administrative data on employment levels and wages and do not find a positive treatment effect with survey data that then disappears with administrative data. In addition, the authors test whether treated firms were more likely to a) provide alternative contacts on the survey, and b) report sales on the follow-up survey, which should be the case if they wanted to please the interviewer. They find no statistically significant differences in either measure across the treatment and control groups, and interpret these findings as indication that self-reporting bias is likely absent. Similarly, Cai and Szeidl (2018) provide evidence on the lack of experimenter demand effect in businesses surveys by comparing self-reported sales with actual book value of sales and finding no statistical difference. It is worth noting that both Bruhn et al. (2018) and Cai and Szeidl (2018) involve interventions which, compared to ours, are much more intense, expensive, and lengthy, which would make their studies more prone to experimenter demand effects.

Second, we have more specific reasons to argue that self-reporting bias is unlikely in our setting as we do not explicitly link the survey to the intervention as in other studies. In addition, if the results were influenced by experimenter demand effects, we would likely observe significant effects for the handbook treatment as well, which we do not. Also, such misreporting would affect all business practices and not just the ones included in the handbook. As a direct test, we run placebo regressions to assess treatment effects on practices that were not mentioned in the handbook. Appendix Table H2 confirms that our treatments did not have any significant impact on these placebo practices, which supports the argument that the impacts on business practice adoption are legitimate effects.

Finally, we use data on "objective" measures gathered by the survey team. At the end of

[27] As highlighted in Carpena and Zia (2020) and Imai et al. (2010), there are two important assumptions needed to identify the ACME and ADE. First is that treatment assignment is independent of the outcomes, which is satisfied in our study because treatment status is randomly assigned. Second, conditional on treatment status, the mediator is also independent of the outcomes. For further discussion on the merits of mediation analysis performed in this paper, see Appendix F.

the baseline and endline surveys, enumerators were instructed to report their own views on shop appearance. For instance, they were asked to record whether the shop appeared clean, well stocked, and whether prices were clearly marked. While these measures do not directly indicate or confirm the use of specific business practices, they do provide a rough indication of whether the shop is better "organized". As Appendix Table H4 shows, we do see improvements on some of these measures in shop owners assigned to *Movie* and *Counseling*. Even though shop owners assigned to *All Three* do not show significant changes in these measures, the sign of the coefficients go in the right direction.

B Unpacking the Counseling Channel: Which Behavioral Constraints are Likely Binding?

The strong and sustained effects of *Counseling* and *All Three* in our study suggest that constraints like procrastination, status quo, and/or inattention may be important impediments to practice and performance improvement. At face value, this is not surprising as people have sticky habits which are difficult to influence (Kahneman et al., 1991), and changing practices that have been implemented for years (sometimes through generations) require time investment with uncertain returns (Verhoogen, 2020). Furthermore, distrust regarding the veracity of information can further restrict behavior change.

The handbook coupled with the counseling in either *Counseling* or *All Three* addresses each of these constraints. It gives retailers a nudge to overcome their procrastination. It also provides credible information from successful peers so the information comes across as accessible and achievable. It links practices to business performance, so the uncertainty about the benefits of investing in changing a habit is reduced. Overall, the handbook with counseling alleviates both information and behavioral constraints.

While it is difficult to isolate how precisely the counseling channel operates, we have some data on the facilitation visits that can shed light on mechanisms. First, we have data on facilitator gender. Since facilitators were randomly assigned to beneficiaries, we have exogenous variation in whether a beneficiary received counseling from a facilitator of the same gender (low social distance) or opposite gender (high social distance). While an imperfect measure of social distance, we believe gender match does carry some weight in a conservative Muslim society such as Indonesia, where it is likely that the level of comfort and acceptance would be higher with matched genders. We use this variation to assess whether there are differences in practice adoption for the *Counseling* and *All Three* treatment groups based on whether there was a gender match between the facilitator and beneficiary. Appendix Table H1 presents these results and finds no statistical difference based on gender match, in fact, the point estimates are almost identical. With the earlier stated caveat on a limited social distance measure, this finding suggests that matching demographics between facilitators and beneficiaries was not a main determinant of impact.

Next, as part of the first shop visit, the facilitators recorded whether the beneficiary had started reading the handbook (extensive margin) and whether the beneficiary expressed difficulty in comprehending the material (intensive margin). These data in Appendix Table C2 show that 38% of beneficiaries visited had not yet started reading the handbook, and of the 62% who had started only 6% expressed difficulties in reading comprehension. Furthermore, among the 38% who had not yet started reading, only 1% stated reading comprehension as a binding constraint. In addition, we do not find heterogeneous impacts by education levels in Appendix Table H5. While not conclusive,

collectively these findings suggest that inattention, status-quo, and/or procrastination are likely to be important channels for why many beneficiaries did not even start reading the handbook even though reading comprehension is not a reported constraint. On the intensive margin, it is very difficult to distinguish from the alternative channel that facilitators could tailor the handbook to specific individual needs. We acknowledge that we can only claim that the counseling related arms operate through some combination of these channels. Dissecting individual components in an experimental setting beyond what we can do with observable data remains an interesting and important question for future research.

C Do Treated Businesses Improve Performance at the Expense of the Control Group?

Another concern relates to whether the gains in sales and profits are achieved at the expense of the control group. In order to allay this concern, we present two key pieces of evidence. First, a simple comparison of sales and profits for the control group between baseline and endline shows that trends in both these variables are flat: USD PPP 4,454 and USD PPP 890, respectively at baseline and USD PPP 4,999 and USD PPP 895, respectively at endline. Neither of these performance variables are statistically different from baseline to endline. Given that sales and profits could have changed in the absence of the treatment for macroeconomic or other external reasons, this evidence suggests a flat rather than an increasing or declining growth trajectory.

Second, to detect geographic spillovers in the spirit of Miguel and Kremer (2004), we use GPS data to directly measure linear distance from each shop in the control group to its nearest shop in the treatment group. If treated shops were improving performance at the expense of others, then control group shops closer to a treated shop would have worse performance outcomes than those located further away. Moreover, we would expect a positive and significant coefficient on the "distance to nearest treated shop" variable in a specification restricted to the control group. Appendix Table H3 presents a simple regression of profits and sales on this distance measure. Although treated shops are randomly assigned and hence location is exogenous, we nevertheless include a control for "distance to nearest control shop" in these regressions to account for market density. The coefficients for both profits and sales are in fact negative and not statistically significant, which suggests that the gains in performance were not achieved at the expense of the control group.[28]

VI Conclusion

This paper shows that it is possible to stimulate the efficiency and growth of small firms by harnessing and curating locally relevant business knowledge. Our findings show that the delivery mechanism of this information is critical – pure information alone, even when carefully curated to cater to the local context, does not have an impact on the adoption of business practices or on performance, but combining it with inexpensive behavioral nudges results in sizable and significant improvements on both dimensions. These effects persist up to 18 months after the interventions, which is indicative of their long-term durability.

[28] Note that we do not have data on all firms in the study area so cannot rule out general equilibrium effects elsewhere. As in most field studies of this type, we acknowledge that our results represent a partial equilibrium.

On scalability, a back-of-the-envelope tabulation reveals the per-firm cost of the *Handbook* intervention was approximately USD 100, with the *Movie* and *Counseling* costing an additional USD 25 each.[29] In contrast, the benefits were up to USD 310 per month in profits, along with a high adoption rate of efficient business practices. This simple comparison shows that the interventions were highly cost-effective with rich potential for scaling up.

The cost effectiveness and scalability are further supported by the local relevance and ease of implementation of our study design. Moreover, the interventions studied in this paper do not require extensive training programs, long business consulting or peer-to-peer sessions, or other heavy time demands from participants. Indeed, other methods studied in the literature that involve face-to-face meetings cost substantially more money, demand higher engagement, and require highly skilled instructors.

Finally, in light of the existing literature on small business growth in developing economies, the overarching insight of this paper is that we are neither in a world where only first principles matter, because otherwise curated information alone would be very effective, nor are we in a world where the keys to business growth are so idiosyncratic that professional consultants or other forms of deeper engagement are required to unlock success. Moreover, our findings suggest that locally useful knowledge is context dependent but not atomistic, and our methodology offers a simple, effective, and inexpensive blueprint for harnessing its potential.

[29] These cost estimates do not include our own research time, but they include the cost of qualitative and quantitative interviews, handbook production, printing, and distribution. The *Movie* and *Counseling* cost estimates additionally include production expenses, facilitator training and salaries, equipment and location rental, and respondent travel and meal compensation. Further, embedded within the cost of interviews and salaries are costs associated with RA time as well as overhead charged by our implementation partner (JPAL) for research services. These include several rounds of piloting and fixed costs associated with getting a new research project off the ground. When scaled, these costs are likely to go down.

References

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Anderson, S., Chandy, R., and Zia, B. (2018). Pathways to Profits: Identifying Separate Channels of Small Firm Growth through Business Training. *Management Science*, 64(12):1–25.

Anderson, S., Lazicky, C., and Zia, B. (2019). Measuring the Unmeasured: Combining Technology and Behavioral Insights to Improve Measurement of Business Outcomes. *World Bank Policy Research Working Paper No. 8836*.

Anderson, S. and McKenzie, D. (2020). Improving Business Practices and the Boundary of the Entrepreneur: A Randomized Experiment Comparing Training, Consulting, Insourcing and Outsourcing. *Mimeo*.

Arrow, K. (1969). Classificatory Notes on the Production and Transmission of Technological Knowledge. *American Economic Review*, 59(2):29–35.

Barsoum, G., Crepon, B., Gardiner, D., Michel, B., and Pariente, W. (2018). Evaluating the Impact of Entrepreneurship Edutainment in Egypt: An experimental approach with peer effects. *Mimeo*.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300.

Benjamini, Y., Krieger, A. M., and Yekutieli, D. (2006). Adaptive Linear Step-up Procedures that Control the False Discovery Rate. *Biometrika*, 93(3):491–507.

Berg, G. and Zia, B. (2017). Harnessing Emotional Connections to Improve Financial Decisions: Evaluating the Impact of Financial Education in Mainstream Media. *Journal of the European Economic Association*, 15(5):1025–1055.

Bernard, T., Dercon, S., Orkin, K., and Taffesse, A. S. (2014). The Future in Mind: Aspirations and Forward-looking Behaviour in Rural Ethiopia. *CEPR Discussion Paper*, 10224.

Bjorvatn, K., Cappelen, A. W., Sekei, L. H., and Sørensen Erik Øand Tungodden, B. (2020). Teaching Through Television: Experimental Evidence on Entrepreneurship Education in Tanzania. *Management Science*, 66(6):2308–2325.

Bloom, N., Brynjolfsson, E., Foster, L., Jarmin, R., Patnaik, M., Saporta-Eksten, I., and Van Reenen, J. (2019). What Drives Differences in Management Practices? *American Economic Review*, 109(5):1648–83.

Bloom, N. and Van Reenen, J. (2007). Measuring and Explaining Management Practices Across Firms and Countries. *The Quarterly Journal of Economics*, 122(4):1351–1408.

Brooks, W., Donovan, K., and Johnson, T. R. (2018). Mentors or Teachers? Microenterprise Training in Kenya. *American Economic Journal: Applied Economics*, 10(4):196–221.

Bruhn, M., Karlan, D., and Schoar, A. (2018). The Impact of Consulting Services on Small and Medium Enterprises: Evidence from a Randomized Trial in Mexico. *Journal of Political Economy*, 126(2):635–687.

Bruhn, M. and McKenzie, D. (2009). In Pursuit of Balance: Randomisation in Practice in Development Field Experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Cai, J. and Szeidl, A. (2018). Interfirm Relationships and Business Performance. *Quarterly Journal of Economics*, 7(2):125–161.

Calderón, G., Cunha, J. M., and De Giorgi, G. (2013). Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico. *Barcelona GSE Working Paper Series*, 742.

Carpena, F. and Zia, B. (2020). The Causal Mechanism of Financial Education: Evidence from Mediation Analysis. *Journal of Economic Behavior and Organization*, 177:143–184.

Chong, A. and La Ferrara, E. (2009). Television and Divorce: Evidence from Brazilian Novelas. *Journal of the European Economic Association*, 7(2-3):458–468.

Chudek, M., Brosseau-Liard, P., Birch, S., and Henrich, J. (2013). Culture-Gene Coevolutionary Theory and Children's Selective Social Learning. In Banaji, M. and Gelman, S., editors, *Navigating the Social World: What Infants, Children, and Other Species Can Teach Us*, pages 1–10. Oxford University Press, Oxford, UK.

Corriveau, K. H. and Harris, P. L. (2009). Choosing Your Informant: Weighing Familiarity and Recent Accuracy. *Developmental Science*, 12(3):426–437.

Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and Aspirations Failure. *Economic Journal*, 126(590):165–188.

de Mel, S., McKenzie, D., and Woodruff, C. (2009). Measuring Microenterprise Profits: Must we Ask How the Sausage is Made? *Journal of Development Economics*, 88(1):19–31.

de Mel, S., Mckenzie, D., and Woodruff, C. (2012). One-Time Transfers of Cash or Capital Have Long-Lasting Effects on Microenterprises in Sri Lanka. *Science*, 335:962–966.

Duflo, E., Kremer, M., and Robinson, J. (2011). Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya. *American Economic Review*, 101(6):2350–90.

Efferson, C., Lalive, R., and Fehr, E. (2008). The Coevolution of Cultural Groups and Ingroup Favoritism. *Science*, 321:1844–1849.

Fafchamps, M. and Quinn, S. (2018). Networks and Manufacturing Firms in Africa: Results from a Randomized Field Experiment. *The World Bank Economic Review*, 32(3):656–675.

Gibbons, R. (2010). Inside Organizations: Pricing, Politics, and Path Dependence. *Annual Review of Economics*, 2(1):337–365.

Gibbons, R. and Henderson, R. (2013). *What Do Managers Do? Exploring Persistent Performance Differences among Seemingly Similar Enterprises*, chapter 17, pages 680–731. The Handbook of Organizational Economics.

Hardy Morgan and McCasland, J. (2016). It Takes Two: Experimental Evidence on the Determinants of Technology Diffusion. *Mimeo*.

Henrich, J. and Gil-White, F. J. (2001). The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission. *Evolution and Human Behavior*, 22:165–196.

Hsieh, C. T. and Olken, B. A. (2014). The Missing 'Missing Middle'. *Journal of Economic Perspectives*, 28(3):89–108.

Imai, K., Keele, L., Tingley, D., and Yamamoto, T. (2011). Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. *American Political Science Review*, 105(4):765–789.

Imai, K., Keele, L., and Yamamoto, T. (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, 25(1):51–71.

Indonesian Ministry of Cooperatives and SMEs Indonesia (2011). Statistik Usaha Mikro, Kecil dan Menengah. Technical report, Kementerian Koperasi dan UMKM.

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., and Puranen, B., editors (2014). *World Values Survey: Round Six*. JD Systems Institute, Madrid.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *Journal of Economic Perspectives*, 5(1):193–206.

King, A. C., Friedman, R., Marcus, B., Castro, C., Napolitano, M., Ahn, D., and Baker, L. (2007). Ongoing Physical Activity Advice by Humans versus Computers: the Community Health Advice by Telephone (CHAT) Trial. *Health psychology: official journal of the Division of Health Psychology, American Psychological Association*, 26:718–27.

La Ferrara, E. (2019). Presidential Address: Aspirations, Social Norms, and Development. *Journal of the European Economic Association*, 17(6):1687–1722.

La Ferrara, E., Chong, A., and Duryea, S. (2012). Soap Operas and Fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, 4(4):1–31.

Lafortune, J., Riutort, J., and Tessada, J. (2018). Role Models or Individual Consulting: The Impact of Personalizing Micro-entrepreneurship Training. *American Economic Journal: Applied Economics*, 10(4):222–245.

Maffioli, A., McKenzie, D., and Ubfal, D. (2020). Estimating the Demand for Business Training: Evidence from Jamaica. *Policy Research Working Paper;No. 9415. World Bank, Washington, DC.*

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., and Waring, T. (2008). Beyond Existence and Aiming Outside the Laboratory: Estimating Frequency-dependent and Pay-off-biased Social Learning Strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363:3515–3528.

McKenzie, D. (2017). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. *American Economic Review*, 107(8):2278–2307.

McKenzie, D. (2020). Small Business Training to Improve Management Practices in Developing Countries: Reassessing the Evidence for 'Training Doesn't Work'. *Policy Research Working Paper No. WPS 9408, Washington, D.C.: World Bank Group.*

McKenzie, D. and Sansone, D. (2019). Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria. *Journal of Development Economics*, 141:102369.

McKenzie, D. and Woodruff, C. (2017). Business Practices in Small Firms in Developing Countries. *Management Science*, 63(9):2967–2981.

Miguel, E. and Kremer, M. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, 72(1):159–217.

Polanyi, M. (1958). *Personal Knowledge: Towards a Post-Critical Philosophy.* University of Chicago Press.

Quinn, S. and Woodruff, C. (2019). Experiments and Entrepreneurship in Developing Countries. *Annual Review of Economics*, 11(1):225–248.

Rendell, L. E., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., and Laland, K. N. (2011). Cognitive Culture: Theoretical and Empirical Insights into Social Learning Strategies. *Trends in Cognitive Sciences*, 15(2):68–76.

Riley, E. (2017). Increasing Students' Aspirations: The Impact of Queen of Katwe on Students' Educational Attainment. *CSAE Working Paper Series*, 2017-13.

Sagiv, L., Arieli, S., Goldenberg, J., and Goldschmidt, A. (2010). Structure and Freedom in Creativity: The Interplay Between Externally Imposed Structure and Personal Cognitive Style. *Journal of Organizational Behavior*, 31:1086–1110.

Statistics Indonesia (2016). Statistical Yearbook of Indonesia 2016. Technical report, Badan Pusat Statistik.

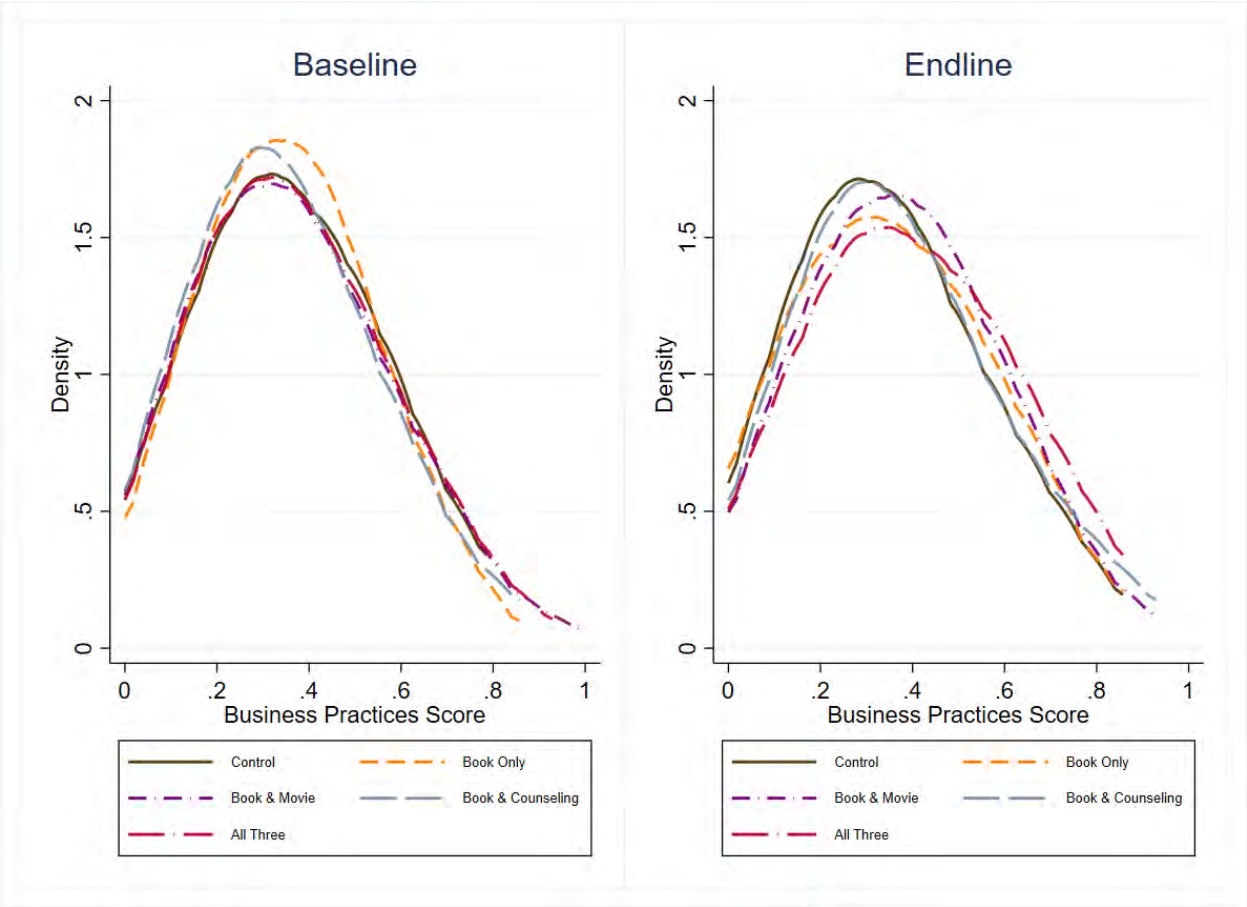
Steiner, J., Stewart, C., and Matéjka, F. (2017). Rational Inattention Dynamics: Inertia and Delay in Decision-Making. *Econometrica*, 85(2):521–553.

Syverson, C. (2011). What Determines Productivity? *Journal of Economic Literature*, 49(2):326–65.

Verhoogen, E. (2020). Firm-Level Upgrading in Developing Countries. *Journal of Economic Literature, forthcoming.*

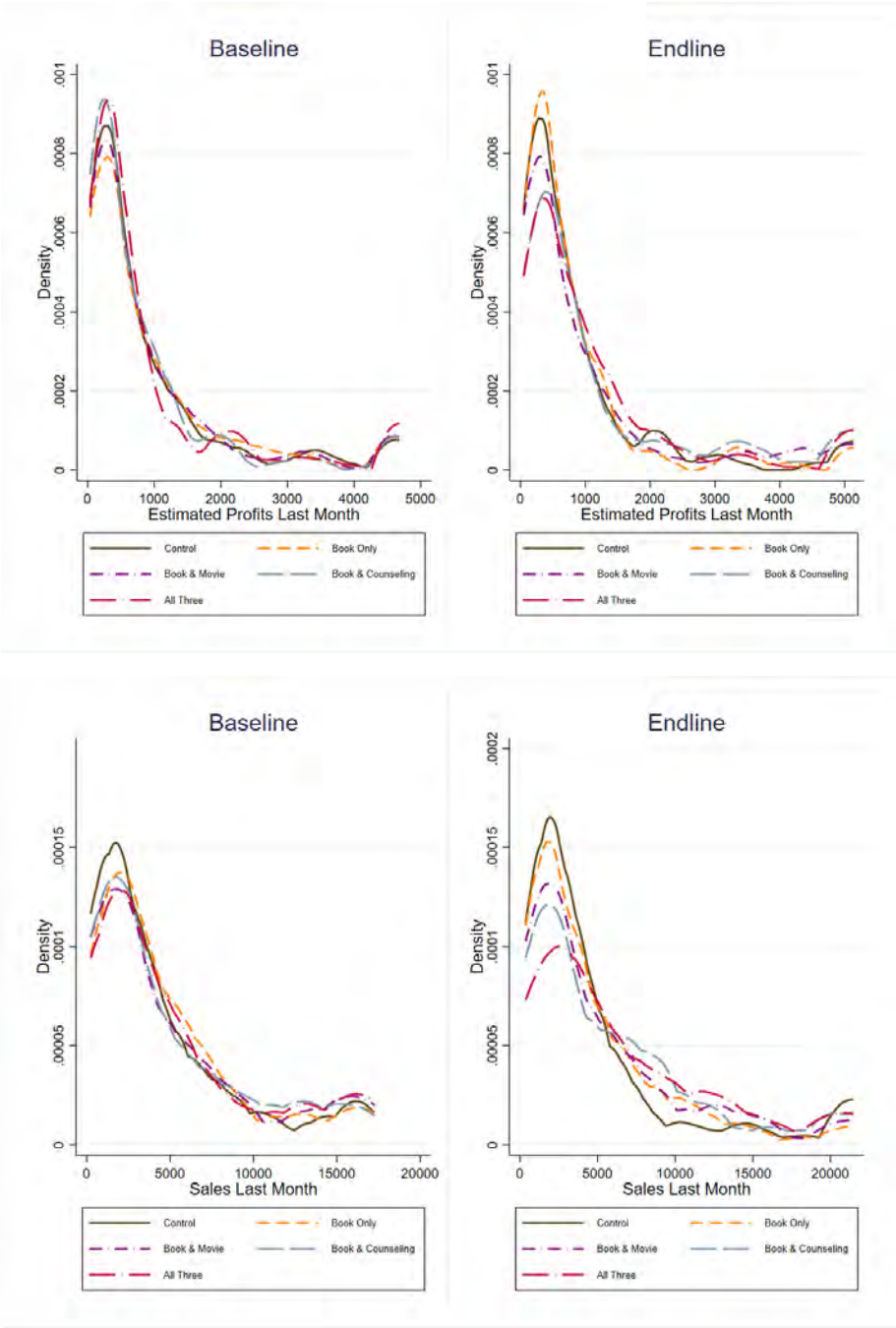
Winter, S. G. (2006). Toward a neo-Schumpeterian theory of the firm. *Industrial and Corporate Change*, 15(1):125–141.

Figure 1: Probability Density Functions of Business Practices Aggregate Score



This figure plots the Probability Density Function (PDF) of the business practices aggregate score at baseline and endline. The figure overlays the PDF of each treatment group and the control group. A Kolmogorov-Smirnov test of equality of distributions cannot be rejected for any combination at baseline. At endline, equality between the "All Three" treatment group and the control group is rejected at the 5% significance level.

Figure 2: Probability Density Functions of Profits and Sales



This figure plots the Probability Density Functions (PDFs) of estimated profits and sales at baseline and endline. For both outcomes, the figure overlays the PDF of each treatment group and the control group. A Kolmogorov-Smirnov test of equality of distributions cannot be rejected for any combination at baseline for either estimated profits or sales. At endline, equality between the "All Three" treatment group and the control group is rejected at the 5% significance level for estimated profits and 1% level for sales. In addition, for estimated profits at endline, equality of distributions is rejected between: (a) "All Three" and "Book Only" at the 5% level; (b) "All Three" and "Book & Movie" at the 10% level; (c) and "Book & Counseling" and "Book Only" at the 10% level. For sales at endline, equality of distributions is rejected between: (a) "All Three" and "Book Only" at the 5% level; (b) "Book & Counseling" and the control group at the 1% level; (c) and "Book & Counseling" and "Book Only" at the 10% level.

Table 1: Summary Statistics and Tests of Randomization

	(1)	(2)	(3)
Variable	Sample Mean	Standard Deviation	F-Test of Joint Equality of Means Across All Treatments (p-value)
Respondent is Male	0.29		0.96
Respondent Age	45.27	11.31	1.00
Respondent Years of Education	9.39	3.78	0.66
Respondent Risk Preference Score (0-1 scale)	0.37	0.21	0.64
Respondent Time Preference Score (0-1 scale)	0.52	0.23	0.92
Shop Size in Square Meters	13.22	12.34	0.63
Age of Firm	13.60	11.79	0.72
Firm has Tax ID	0.19		0.37
Total Number of Employees	2.00	1.22	0.58
Total Number of Customers on a Normal Day	49.33	43.32	0.29
Sales Last Month (USD PPP)	4786.16	4853.48	0.63
Estimated Profits Last Month (USD PPP)	907.73	1185.09	0.71
Obtained Business Loan in Last 12 Months	0.16		0.52
Aggregate Score of Practices Covered in Handbook (0-1 scale)	0.36	0.18	0.83
Aggregate Score for Practices Not Covered in Handbook (0-1 scale)	0.18	0.13	0.59
Aggregate Score of Enumerator Shop Evaluation (0-1 scale)	0.63	0.24	0.52

Omnibus Chi-Square Test of Joint Orthogonality from Multinomial Logit (p-value): 0.819

Notes: This table presents summary statistics for the baseline survey data. Columns 1 and 2 present mean and standard deviations for the full sample. Column 3 presents p-values for equality of means tests across all experimental groups (treatment and control). The bottom of the table presents the p-value from a chi-square test of joint orthogonality of all variables reported in the table using a multinomial logit specification.

Table 2: Business Practices Domain

	(1)	(2)	(3)	(4)	(5)	(6)
	Aggregate of All Practices	Record Keeping	Planning	Stocking Up	Marketing	Joint Decision Making
Assigned Handbook	0.008	0.025	0.028	-0.007	-0.010	0.011
	(0.013)	(0.017)	(0.022)	(0.019)	(0.022)	(0.027)
	0.157	*0.127*	*0.167*	*0.290*	*0.290*	*0.290*
Assigned Handbook & Movie	0.042***	0.058***	0.044*	0.038*	0.042	0.039
	(0.013)	(0.018)	(0.021)	(0.020)	(0.023)	(0.028)
	0.002	*0.009*	*0.071*	*0.086*	*0.103*	*0.142*
Assigned Handbook & Counseling	0.034***	0.065***	0.034	0.011	0.039	0.037
	(0.013)	(0.018)	(0.021)	(0.019)	(0.023)	(0.027)
	0.006	*0.005*	*0.111*	*0.290*	*0.111*	*0.152*
Assigned All Three	0.054***	0.054***	0.068***	0.053**	0.061**	0.059*
	(0.013)	(0.017)	(0.022)	(0.019)	(0.024)	(0.027)
	0.001	*0.009*	*0.009*	*0.017*	*0.026*	*0.070*
R-squared	0.308	0.205	0.193	0.187	0.152	0.120
N	2205	2205	2204	2205	2205	2205
Dependent Variable Mean in Control Group	0.337	0.196	0.402	0.471	0.250	0.269
Dependent Variable SD in Control Group	0.189	0.252	0.310	0.270	0.320	0.420
F-test (p-value): Book = Book & Movie	0.008	0.063	0.464	0.013	0.024	0.305
F-test (p-value): Book = Book & Counseling	0.040	0.026	0.769	0.307	0.032	0.346
F-test (p-value): Book = All Three	0.000	0.104	0.078	0.001	0.002	0.082
F-test (p-value): Book & Movie = Book & Counseling	0.539	0.701	0.629	0.144	0.885	0.928
F-test (p-value): Book & Movie = All Three	0.363	0.826	0.276	0.421	0.451	0.488

Notes: This table presents regression analysis for the business practices domain, including the aggregate of all practices (column 1) and sub-practice indices of record keeping (column 2), planning (column 3), stocking-up (column 4), marketing (column 5), and joint decision making (column 6). The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Sharpened q-values that correct for false discovery rates using the method of Benjamini et al. (2006) are reported in italics. For column (1), FDR corrections are performed across treatment arms. For columns (2)-(6), FDR corrections are simultaneously performed across all tests (i.e. # of treatment arms * # of outcomes). Statistically significant q-values are highlighted by: * (10% significance level), ** (5% significance level), and *** (1% significance level).

Table 3: Business Performance Domain

	(1)	(2)	(3)
	Performance Domain Z-Score	Estimated Profits Last Month (win 5%)	Sales Last Month (win 5%)
Assigned Handbook	-0.038	-90.194	-384.190
	(0.029)	(78.342)	(312.447)
	0.153	*0.162*	*0.162*
Assigned Handbook & Movie	0.030	113.842	365.202
	(0.031)	(86.829)	(336.910)
	0.200	*0.162*	*0.162*
Assigned Handbook & Counseling	0.100**	309.980***	835.637**
	(0.033)	(89.633)	(373.159)
	0.011	*0.005*	*0.047*
Assigned All Three	0.074**	190.267**	803.081**
	(0.032)	(84.814)	(357.874)
	0.034	*0.047*	*0.047*
R-squared	0.399	0.180	0.493
N	2197	2172	2197
Dependent Variable Mean in Control Group	-0.123	894.544	4998.923
Dependent Variable SD in Control Group	0.481	1127.783	5623.257
F-test (p-value): Book = Book & Movie	0.026	0.018	0.017
F-test (p-value): Book = Book & Counseling	0.000	0.000	0.000
F-test (p-value): Book = All Three	0.000	0.001	0.000
F-test (p-value): Book & Movie = Book & Counseling	0.037	0.043	0.203
F-test (p-value): Book & Movie = All Three	0.182	0.409	0.221

Notes: This table presents regression analysis for the business performance domain, including the average z-score of profits and sales (column 1) as well as the individual outcomes of estimated profits (column 2) and sales (column 3). The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Sharpened q-values that correct for false discovery rates using the method of Benjamini et al. (2006) are reported in italics. For column (1), FDR corrections are performed across treatment arms. For columns (2)-(3), FDR corrections are simultaneously performed across all tests (i.e. # of treatment arms * # of outcomes). Statistically significant q-values are highlighted by: ** (5% significance level), and *** (1% significance level).

Table 4: Business Expenditure Domain

	(1)	(2)	(3)	(4)	(5)	(6)
	Total Expenses Last Month (win 5%)	Stocking Up (win 5%)	Wages (win 5%)	Rent (win 5%)	Electricity (win 5%)	Transport (win 5%)
Assigned Handbook	-110.180	-97.514	-2.385	-6.854	1.957	0.348
	(253.384)	(249.141)	(1.593)	(5.005)	(3.508)	(2.199)
	1.000	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*
Assigned Handbook & Movie	154.728	128.388	-1.293	-1.342	3.266	0.074
	(283.544)	(278.493)	(1.905)	(4.893)	(3.397)	(2.231)
	1.000	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*
Assigned Handbook & Counseling	174.898	188.687	2.784	-1.271	5.955*	-0.048
	(294.527)	(284.571)	(2.120)	(4.986)	(3.394)	(2.112)
	1.000	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*
Assigned All Three	334.029	321.822	-1.620	-3.085	-0.685	-0.342
	(290.281)	(287.664)	(1.601)	(4.763)	(3.373)	(2.156)
	1.000	*1.000*	*1.000*	*1.000*	*1.000*	*1.000*
R-squared	0.532	0.535	0.314	0.628	0.285	0.223
N	2188	2183	2187	2187	2172	2180
Dependent Variable Mean in Control Group	4287.789	4093.664	7.072	51.573	66.853	30.601
Dependent Variable SD in Control Group	4811.178	4702.590	30.335	90.935	50.646	29.975
F-test (p-value): Book = Book & Movie	0.303	0.367	0.582	0.220	0.717	0.901
F-test (p-value): Book = Book & Counseling	0.294	0.271	0.017	0.222	0.275	0.849
F-test (p-value): Book = All Three	0.097	0.111	0.646	0.380	0.467	0.743
F-test (p-value): Book & Movie = Book & Counseling	0.946	0.833	0.093	0.987	0.454	0.954
F-test (p-value): Book & Movie = All Three	0.542	0.504	0.868	0.678	0.264	0.846

Notes: This table presents regression analysis for the business expenditure domain, including the total expenses (column 1) as well as the individual expense categories of stocking up (column 2), wages (column 3), rent (column 4), electricity (column 5), and transport (column 6). The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Sharpened q-values that correct for false discovery rates using the method of Benjamini et al. (2006) are reported in italics. For column (1), FDR corrections are performed across treatment arms. For columns (2)-(6), FDR corrections are simultaneously performed across all tests (i.e. # of treatment arms * # of outcomes).

Table 5: Mechanisms Domain

	(1)	(2)	(3)	(4)
	Mechanism Domain Z-Score	Shop Size in Square Meters	Total Number of Employees	Total Number of Customers
Assigned Handbook	-0.007	-0.129	0.013	-0.351
	(0.038)	(0.666)	(0.086)	(2.452)
	1.000	*1.000*	*1.000*	*1.000*
Assigned Handbook & Movie	0.035	0.132	0.076	1.077
	(0.039)	(0.676)	(0.083)	(2.212)
	1.000	*1.000*	*1.000*	*1.000*
Assigned Handbook & Counseling	0.066	0.934	0.130	0.635
	(0.041)	(0.716)	(0.079)	(2.421)
	0.764	*1.000*	*1.000*	*1.000*
Assigned All Three	0.021	0.388	0.010	1.456
	(0.038)	(0.635)	(0.084)	(2.414)
	1.000	*1.000*	*1.000*	*1.000*
R-squared	0.365	0.357	0.227	0.336
N	2205	2204	2205	2203
Dependent Variable Mean in Control Group	-0.049	12.856	1.954	40.232
Dependent Variable SD in Control Group	0.551	9.235	1.150	35.578
F-test (p-value): Book = Book & Movie	0.297	0.702	0.468	0.573
F-test (p-value): Book = Book & Counseling	0.084	0.135	0.167	0.724
F-test (p-value): Book = All Three	0.465	0.414	0.976	0.511
F-test (p-value): Book & Movie = Book & Counseling	0.475	0.258	0.513	0.862
F-test (p-value): Book & Movie = All Three	0.732	0.685	0.442	0.877

Notes: This table presents regression analysis for the mechanisms domain, including the average z-score of shop size, employees, and customers (column 1); as well as the individual outcomes of shop size (column 2), number of employees (column 3), and number of customers (column 4). The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Sharpened q-values that correct for false discovery rates using the method of Benjamini et al. (2006) are reported in italics. For column (1), FDR corrections are performed across treatment arms. For columns (2)-(4), FDR corrections are simultaneously performed across all tests (i.e. # of treatment arms * # of outcomes).

Table 6: Personal Attributes Domain

	(1)	(2)	(3)	(4)	(5)
	Satisfied with Life (0-1 scale)	Satisfied with Finances (0-1 scale)	Systematic Working Style (0-1 scale)	Intuitive Working Style (0-1 scale)	Aspirations Z-score Index
Assigned Handbook	0.031	0.022	-0.034	-0.015	-0.045
	(0.021)	(0.021)	(0.016)	(0.015)	(0.025)
	1.000	*1.000*	*0.409*	*1.000*	*0.819*
Assigned Handbook & Movie	0.000	0.022	-0.009	0.000	-0.003
	(0.021)	(0.021)	(0.015)	(0.015)	(0.026)
	1.000	*1.000*	*1.000*	*1.000*	*1.000*
Assigned Handbook & Counseling	0.016	0.020	-0.002	-0.024	0.024
	(0.020)	(0.020)	(0.015)	(0.016)	(0.026)
	1.000	*1.000*	*1.000*	*1.000*	*1.000*
Assigned All Three	0.015	0.045	-0.019	-0.004	0.002
	(0.019)	(0.020)	(0.015)	(0.016)	(0.026)
	1.000	*0.409*	*1.000*	*1.000*	*1.000*
R-squared	0.040	0.037	0.158	0.101	0.357
N	1019	1018	1181	1181	2205
Dependent Variable Mean in Control Group	0.699	0.635	0.799	0.618	-0.091
Dependent Variable SD in Control Group	0.212	0.211	0.182	0.171	0.379
F-test (p-value): Book = Book & Movie	0.152	0.969	0.091	0.309	0.107
F-test (p-value): Book = Book & Counseling	0.474	0.920	0.039	0.573	0.009
F-test (p-value): Book = All Three	0.414	0.256	0.307	0.477	0.074
F-test (p-value): Book & Movie = Book & Counseling	0.455	0.887	0.656	0.133	0.322
F-test (p-value): Book & Movie = All Three	0.464	0.271	0.469	0.781	0.877

Notes: This table presents regression analysis for the personal attributes domain, including a measure of the business owner's satisfaction with life (column 1), a measure of satisfaction with finances (column 2), a measure of systemic working style (column 3), a measure of intuitive working style (column 4), and an aspirations z-score index (column 5). The aspirations index is the average z-score of 12-month aspirations for shop size, employment, customers, and sales. Columns (1) and (2) outcomes are only available for the endline; and columns (3) and (4) outcomes are only available for the baseline and midline. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy where applicable, baseline value of the outcome variable where applicable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Sharpened q-values that correct for false discovery rates using the method of Benjamini et al. (2006) are reported in italics. FDR corrections are simultaneously performed across all tests (i.e. # of treatment arms * # of outcomes) reported in the table.

Online Appendix

Curating Local Knowledge:
Experimental Evidence from Small Retailers in Indonesia

Appendix A: Comparing the Handbook to Other Approaches

This appendix compares the handbook used in this study to other common approaches used in research and practice.

Handbook vs Standardized Classroom Business Training

There are several differences between the handbook used in this study and a standard business training program. The most common and well-established formal business training course is the International Labor Organization's (ILO) Start and Improve your Business (SIYB) program, which has an estimated outreach of 6 million trainees and a network of over 17,000 trainers over the last 30 years.[1]

The SIYB implementation plan involves capacity building through comprehensive training-of-trainer modules and certification of master trainers. The average cost of these training ranges between 50,000 USD -- 150,000 USD, and the master trainer development process can take up to 2 years. Moreover, these costs do not include delivery to beneficiaries; rather, they are the fixed upfront costs of building capacity of the trainers. Given that the focus of the SIYB program is on providing classroom training and on-site individual consulting, the marginal costs for beneficiary training are substantial. Furthermore, prior research has shown that interest and attendance in classroom training can be exceedingly low (see for example, Bruhn et al., 2014). By contrast, our handbook relies exclusively on self-learning with no classroom training involved, and the behavioral interventions discussed in the main paper offer very light touch add-ons to the handbook. As such, the cost of delivery is comparatively very low.

While the SIYB program covers some of the same aggregate dimensions of business operations, such as record-keeping, stocking-up, marketing, and planning, the main difference with the handbook is that the latter provides locally relevant ordering of business practices within these broader categories. For instance, while the SIYB syllabus emphasizes all business practices associated with record-keeping as being important, the handbook selects and highlights only those record-keeping practices that have been shown (through the qualitative and quantitative process) to be profitable and locally relevant. Hence, rather than overwhelming business owners with a plethora of business practices that they should adopt, the handbook offers streamlined information on the most profitable practices for their local context. As an illustration of the difference in terms of amount of information provided, the handbook is 60 pages long while the SIYB manual has more than 600 pages. In addition, the handbook offers practical implementation norms and tips derived from the experiences of successful peers and does so by using language and terminology that is native to the local context. By contrast, the SIYB curriculum is highly focused on standardized content and professional language, which may lead to a large disconnect with many beneficiaries.

[1] Information about the SIYB program is available here: SIYB.

Handbook vs Rules-of-Thumb Training

Drexler et al. (2014) propose a rules-of-thumb training that teaches basic financial heuristics to microentrepreneurs and compare this training to a standard accounting training. To develop the rules-of-thumb curriculum, the authors worked with local experts. In that sense, they also use local knowledge to come up with a set of simple practices. However, there are fundamental differences between their approach and ours. First, the local knowledge they use comes from the subjective experiences of the local expert instead of being derived from detailed and structured qualitative and quantitative surveys of firms themselves. In that sense, we circumvent the risk of not finding the right local expert, which is costly and difficult to verify. Second, our approach provides a more objective and quantifiable identification of local best practices as it does not rely on the views of one expert, rather a representative sample of local firms. Third, the rules-of-thumb intervention was delivered as a standard classroom training (over 15 hours). Our intervention does not involve classrooms. This feature of our study is very important not only because it substantially reduces beneficiary opportunity and transport costs, but also because it increases treatment compliance and, hence, statistical power.

Handbook vs Market Intelligence

While our approach has similarities with gathering market intelligence, there are some distinct differences. Foremost, the handbook in our study does not provide insights about the market, such as aggregate trends or information on specific competitors and potential customers of the business. Instead, the emphasis is on the internal practices of a firm. Specifically, our approach involves identifying what works for businesses that share a similar business model, that are of similar size, and share a common culture. While our approach could be extended to represent or mimic market intelligence, this would involve gathering an entirely different set of information from a diverse set of firms and markets.

Appendix A References:

Bruhn, Miriam, Gabriel Lara Ibarra, and David McKenzie. 2014. “The minimal impact of a large-scale financial education program in Mexico City.” *Journal of Development Economics*, 108, 184-189.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar. 2014. “Keeping it simple: Financial literacy and rules of thumb.” *American Economic Journal: Applied Economics*, 6(2), 1-31.

Appendix B: Survey Attrition

There are two sources of survey attrition in our study. Some shops closed down and a few others refused to be part of the midline and endline surveys. Appendix Table B1 below presents regression analysis on survey attrition. Columns (1) and (2) study differential attrition based on refusals both at midline and endline, respectively. Columns (3) and (4) repeat the analysis for attrition based on the shop being closed, also for midline and endline, respectively.

Overall, attrition is very low. We were able to reach 92% of the sample at midline and 81% of the sample at endline, and the small attrition rates are not correlated with treatment status. Specifically, columns (1) and (2) show that there was no differential attrition by treatment, as all the coefficients of the treatment indicators are close to zero and not statistically significant. The magnitude of the coefficients is also small, with 1-4% differences in attrition rate between the treatment arms and the control group. Columns (3) and (4) further show that most businesses remained open through the study period, with only 5.4% of businesses closed by midline and 12.6% by endline; and that this movement on the extensive margin is uncorrelated with any treatment status.

Appendix Table B2 presents Lee bounds for the main practices aggregate and the business performance z-score, comparing unconditional means of trimmed subsamples for all treatment arms against the control group. The proportion of the sample trimmed is small across all treatment arms given the findings of Appendix Table B1, and the “All Three” group shows statistically significant impacts on both outcomes even at the lower bound. The lower bound for “Book & Assistance” is also significant for the performance z-score. These results show that the main treatment effects of this study are robust to concerns related to attrition.

Appendix Table B1: Attrition Analysis

	(1)	(2)	(3)	(4)
	Completed Midline Survey	Completed Endline Survey	Business Closed at Midline Survey	Business Closed at Endline Survey
Assigned Handbook	-0.021	-0.035	0.010	0.004
	(0.024)	(0.035)	(0.020)	(0.029)
Assigned Handbook & Movie	-0.026	-0.007	-0.005	-0.003
	(0.024)	(0.035)	(0.019)	(0.029)
Assigned Handbook & Counseling	-0.022	-0.021	0.009	-0.001
	(0.024)	(0.035)	(0.020)	(0.028)
Assigned All Three	-0.035	-0.033	-0.005	-0.009
	(0.025)	(0.035)	(0.019)	(0.028)
R-squared	0.034	0.043	0.036	0.043
N	1301	1301	1301	1301
Dependent Variable Mean in Control Group	0.927	0.805	0.054	0.126
F-test (p-value): Book = Book & Movie	0.866	0.440	0.463	0.792
F-test (p-value): Book = Book & Assistance	0.985	0.713	0.980	0.853
F-test (p-value): Book = All Three	0.617	0.962	0.465	0.651
F-test (p-value): Book & Movie = Book & Counseling	0.878	0.682	0.471	0.935
F-test (p-value): Book & Movie = All Three	0.744	0.468	0.984	0.854

Notes: This table presents regression analysis for completion of the midline survey (columns 1) and endline survey (columns 2). Columns (3) and (4) repeat the main regression specification for business closures at midline and endline, respectively. All specifications include stratification controls, and use robust standard errors.

Appendix Table B2: Lee Bounds of Impact on Business Practices and Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Aggregate of All Practices				Performance Domain Z-Score			
	Assigned Handbook	Assigned Handbook & Movie	Assigned Handbook & Counseling	Assigned All Three	Assigned Handbook	Assigned Handbook & Movie	Assigned Handbook & Counseling	Assigned All Three
Lee Lower Bound	-0.001	0.037**	0.019	0.044***	-0.038	0.042	0.086**	0.083**
	(0.015)	(0.016)	(0.015)	(0.015)	(0.033)	(0.035)	(0.036)	(0.036)
Lee Upper Bound	0.022	0.053***	0.037**	0.070***	0.052	0.106**	0.168***	0.177***
	(0.016)	(0.017)	(0.017)	(0.017)	(0.049)	(0.050)	(0.051)	(0.053)
N	1042	1042	1042	1042	1042	1042	1042	1042
Proportion Trimmed	.0295	.0206	.0228	.0339	.0317	.0228	.0295	.0406

Notes: This table presents Lee Bounds for the impact on the aggregate of business practices (columns 1-4) and the performance domain z-score (columns 5-8). The sample is pooled across the two follow-up survey rounds, and a round dummy is included in the Lee Bounds estimation. Unconditional means of trimmed subsamples are compared across treatment arms in each column.

Appendix C: Treatment Compliance, Evaluation, and Selection

This appendix discusses treatment compliance and attendee evaluation, summarizes statistics collected during the counseling visits, and presents analysis for selection into treatment.

Treatment Compliance and Evaluation

Appendix Table C1 below presents statistics for treatment compliance and attendee evaluation for the documentary movie and counseling interventions. Panel A focuses on the movie and panel B on the counseling visits.

Panel A shows that of the 520 shop owners invited to the movie screening in the "Handbook and Movie" (column 1) and "All Three" (column 2) treatment group, 52% and 49%, respectively, showed up at the venue. Column 3 does not find differential attendance across these two treatment arms. These compliance rates are in line with previous experience of low takeup for interventions requiring attendance. For example, Drexler et al. (2014), Gine and Mansuri (2014), and Bruhn et al. (2018) document similarly low attendance rates for their respective classroom sessions. Other studies such as Bruhn and Zia (2013) observe even lower attendance, of below 40% of invitees.

Panel A also shows that among those who attended, the feedback from the documentary movie screening was quite positive. Shop owners report having learned something new, and feeling inspired and hopeful after watching the movie. They also did not feel bored during the screening.

Panel B of Appendix Table C1 repeats the above analysis for the counseling intervention. Out of the 520 shop owners offered the counseling sessions, 77% in the "Handbook and Counseling" treatment group and 78% in the "All Three" group accepted the first session. The second session had slightly lower acceptance, at 68% apiece for both treatment groups. This higher participation rate is likely due to the fact that the counseling visits were conducted on site and at a pre-arranged time, thus allowing shop owners to continue monitoring their business activities. The feedback from the counseling sessions was also quite positive, as shown in the remainder of panel B.

Statistics from Counseling Visits

Appendix Table C2 summarizes statistics that were collected by the counselors during their two shop visits. The table shows that at the time of the first visit, 62% of shop owners had started reading the handbook, with some variation in which chapters they read as shown in columns (2)-(6). 60% of shop owners had tried implementing a practice from the handbook, and on a 1-5 scale, they expressed high levels of interest in implementing the handbook practices (average score of 2.96).

Panel B shows that by the second visit, 79% of shop owners had started reading the handbook and 84% had tried implementing a practice.

Treatment Selection

Appendix Table C3 analyzes whether treatment compliance can be explained by observable respondent and business characteristics at baseline. These explanatory variables include the shop owner's gender, age, education level, risk tolerance, and time preference; as well as business size, age, tax status, number of employees, sales, profits, and aggregate practices score. The results across columns for each respective treatment group show that there are no consistent discernable patterns of selection on observables. Shop owners in the handbook only group who accept the handbook are slightly older, run older firms, are more likely to be formal, and have marginally more employees. However, the takeup rate in this group is 90.4% so these differences are very small in real economic terms. Moreover, these selection patterns do not persist across the other treatment groups in columns (2)-(4).

Appendix C References:

Bruhn, Miriam, Dean Karlan, and Antoinette Schoar. 2018. "The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico." *Journal of Political Economy*, 126 (2), 635-687.

Bruhn, Miriam and Bilal Zia. 2013. "Stimulating managerial capital in emerging markets: the impact of business training for young entrepreneurs." *Journal of Development Effectiveness*, 5(2), 232-266.

Calderon, Gabriela, Jesse M. Cunha, and Giacomo De Giorgi. 2013. "Business literacy and development: Evidence from a randomized controlled trial in rural Mexico." *National Bureau of Economic Research,* No. w19740.

Drexler, Alejandro, Greg Fischer, and Antoinette Schoar. 2014. "Keeping it simple: Financial literacy and rules of thumb." *American Economic Journal: Applied Economics*, 6(2), 1-31.

Giné, Xavier, and Ghazala Mansuri. 2014. "Money or Management? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan." *World Bank Working Paper Series*, 6959.

McKenzie, David and Chris Woodruff. 2014. "What are we Learning from Business Training and Entrepreneurship Evaluations around the Developing World?" *World Bank Research Observer*, 29(1): 48-82.

Appendix Table C1: Takeup and Evaluation of Documentary Movie and Counseling

	(1)	(2)	(3)
Panel A: Documentary Movie	Handbook and Movie (A)	All Three (B)	A=B (p-value)
Attended Movie Screening	0.52	0.49	0.54
Evaluation: Learned Something New (0-4 scale)	3.34	3.21	0.18
Evaluation: Feeling Inspired (0-4 scale)	3.31	3.3	0.941
Evaluation: Feeling Hopeful (0-4 scale)	3.6	3.42	0.043**
Evaluation: Feeling Bored (0-4 scale)	0.83	0.97	0.43
Panel B: Counseling	Handbook and Counseling (A)	All Three (B)	A=B (p-value)
Received First Counseling Session	0.77	0.78	0.752
Received Second Counseling Session	0.68	0.68	0.925
Evaluation: Learned Something New (0-4 scale)	2.88	2.89	0.908
Evaluation: Feeling Inspired (0-4 scale)	2.76	2.83	0.422
Evaluation: Feeling Hopeful (0-4 scale)	2.88	2.97	0.312
Evaluation: Feeling Bored (0-4 scale)	0.59	0.43	0.118

Notes: This table presents take-up proportions and evaluations of the documentary movie (Panel A) and counseling (Panel B), separately for the respective treatment arms. The evaluation questions were coded on a 0-4 scale with 0 representing "strongly disagree" and 4 representing "strongly agree." Column (3) presents p-values for equality of means tests for outcomes in columns (1) and (2). Statistically significant p-values are highlighted by: ** (5% significance level).

Appendix Table C2: Statistics from Counseling Visits

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: First Counseling Visit	Full Handbook	Record Keeping	Profit Calculation	Stocking Up	Attracting Customers	Joint Decision Making
Started Reading Content	0.62	0.59	0.48	0.46	0.44	0.40
Tried Implementing Any Practice	0.60	0.40	0.22	0.34	0.49	0.49
Interest in Implementing Practices (1-5 scale)	2.96	2.99	2.99	2.61	2.55	3.73
	(0.29)	(1.52)	(1.33)	(1.26)	(1.35)	(1.31)
Panel B: Second Counseling Visit	Full Handbook	Record Keeping	Profit Calculation	Stocking Up	Attracting Customers	Joint Decision Making
Started Reading Content	0.79	0.77	0.70	0.69	0.67	0.64
Tried Implementing Any Practice	0.84	0.50	0.34	0.44	0.66	0.69

Notes: This table summarizes the data collected by facilitators during the first and second counseling visits. The variables include: whether retailers had started reading the content of the handbook; whether retailers had tried to implement any of the practices described in the handbook; and the level of interest in implementing practices described in the handbook (1-5 scale with 1 representing least interest, and 5 representing most interest). Column (1) presents statistics for the handbook as a whole, and subsequent columns present statistics separately for each chapter.

Appendix Table C3: Selection into Treatment

	(1)	(2)	(3)	(4)
	Received Handbook	Received Handbook & Movie	Received Handbook & Counseling	Received All Three
Respondent is Male	-0.050	0.010	0.025	-0.123
	(0.043)	(0.079)	(0.065)	(0.082)
Respondent Age	0.006***	0.001	-0.002	0.005
	(0.002)	(0.004)	(0.003)	(0.004)
Respondent Years of Education	-0.001	0.013	-0.006	0.010
	(0.006)	(0.010)	(0.008)	(0.010)
Respondent Risk Preference Score (0-1 scale)	0.099	0.077	-0.134	-0.116
	(0.101)	(0.160)	(0.160)	(0.156)
Respondent Time Preference Score (0-1 scale)	-0.029	0.097	-0.001	0.005
	(0.079)	(0.156)	(0.130)	(0.146)
Shop Size (Sq. Meters)	-0.001	0.001	-0.000	-0.003
	(0.002)	(0.003)	(0.003)	(0.002)
Age of Firm	0.002**	-0.001	0.006**	0.004
	(0.001)	(0.003)	(0.003)	(0.003)
Firm has Tax ID	0.240***	0.059	-0.025	0.036
	(0.055)	(0.101)	(0.099)	(0.106)
Total Number of Employees	0.023*	0.010	-0.003	0.032
	(0.012)	(0.030)	(0.029)	(0.023)
Sales Last Month (win 5%)	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Estimated Profits Last Month (win 5%)	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Aggregate Practices Score	0.063	-0.107	0.004	-0.067
	(0.113)	(0.185)	(0.182)	(0.178)
R-squared	0.215	0.186	0.161	0.234
N	260	260	260	260
Dependent Variable Mean	0.904	0.512	0.773	0.469

Notes: This table presents regression analysis on the determinants of take-up of the handbook (column 1), handbook and movie (column 2), handbook and counseling (column 3), and all three (column 4). Statistically significant p-values are highlighted by: *(10% significance level), ** (5% significance level)), and *** (1% significance level).

Appendix D: Business Practices -- Aggregation Robustness, Impacts on Individual Practices, and Quantile Regressions

This appendix presents robustness analysis for the aggregate practices measure used in the main paper in Appendix Table D1. The appendix also provides regressions results for individual business practices in Appendix Tables D2-D6, as well as quantile regression analysis for the aggregate practices measure in Appendix Table D7.

The analysis in the main paper uses an aggregate measure of all business practices that were included in the handbook as a main outcome indicator. This measure is simply the proportion of total practices adopted – by construction this index ranges from 0-1 where 0 indicates that no practices were adopted and 1 indicates that all practices were adopted.

This appendix provides robustness analysis by using three alternative aggregation methods for practices. Appendix Table D1 compares the aggregate of all practices (column 1) with the first principal component of all practices in column (2), as well as two separate lasso specifications that identify the subset of practices that best predict baseline profits (column 3) and sales (column 4), respectively. The post-lasso coefficients are then used to weight the practices and form the two alternative practice composites. We find these three alternative measures are highly correlated with our aggregate index, with correlation coefficients of 0.99, 0.75, and 0.80, respectively.

Moreover, we find very consistent treatment effects across the different columns of Appendix Table D1. Apart from the “Handbook” group, all coefficients in the table are statistically significant after controlling for false discovery. The treatment coefficients cannot be directly compared across measures but can be normalized against the control group standard deviation. We find that within treatment, impacts as a proportion of control group standard deviation are stable across the four weighting methods, ranging between 0.02-0.04 SD for the “Handbook” group; 0.19-0.22 SD for the “Handbook & Movie” group; 0.14-0.19 SD for the “Handbook & Counseling” group; and 0.19-0.29 for the “All Three” group.

These robustness tests show that the positive treatment impacts, effect sizes, and statistical significance do not deviate when we consider alternative weighting schemes.

Appendix Table D1: Business Practices Domain -- Robustness to Alternative Weighting Methods

	(1)	(2)	(3)	(4)
	Aggregate of All Practices	First Principal Component of All Practices	Lasso Baseline Profits	Lasso Baseline Sales
Assigned Handbook	0.008	0.065	2.832	27.530
	(0.013)	(0.111)	(12.460)	(86.265)
	0.190	*0.190*	*0.258*	*0.250*
Assigned Handbook & Movie	0.042***	0.347***	37.948***	234.157***
	(0.013)	(0.112)	(13.133)	(88.353)
	0.006	*0.006*	*0.008*	*0.010*
Assigned Handbook & Counseling	0.034***	0.318***	26.232**	185.474**
	(0.013)	(0.112)	(12.654)	(87.195)
	0.010	*0.008*	*0.020*	*0.019*
Assigned All Three	0.054***	0.469***	35.219***	282.115***
	(0.013)	(0.112)	(13.067)	(89.554)
	0.001	*0.001*	*0.010*	*0.006*
R-squared	0.308	0.311	0.199	0.217
N	2205	2204	2204	2204
Dependent Variable Mean in Control Group	0.337	-0.148	259.815	1830.941
Dependent Variable SD in Control Group	0.189	1.643	183.234	1256.790
F-test (p-value): Book = Book & Movie	0.008	0.012	0.005	0.014
F-test (p-value): Book = Book & Counseling	0.040	0.023	0.050	0.055
F-test (p-value): Book = All Three	0.000	0.000	0.009	0.003
F-test (p-value): Book & Movie = Book & Counseling	0.539	0.793	0.353	0.566
F-test (p-value): Book & Movie = All Three	0.363	0.291	0.835	0.582

Notes: This table presents regression analysis for alternative aggregation methods for business practices. Column (1) is the measure we use in the main paper; column (2) is the first principal component of all practices; and columns (3) and (4) use lasso to choose the subset of practices that best predict baseline profits and sales, respectively, with post-lasso coefficients used as weights for the respective practice measures. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Sharpened q-values that correct for false discovery rates using the method of Benjamini et al. (2006) are reported in italics. FDR corrections are simultaneously performed across all tests (i.e. # of treatment arms * # of outcomes) reported in the table. Statistically significant q-values are highlighted by: ** (5% significance level), and *** (1% significance level).

Appendix Table D2: Impact on Individual Record Keeping Practices

	(1)	(2)	(3)
	Recorded Every Purchase and Sale	Itemized Business Revenues and Expenses	Estimated Cash on Hand
Assigned Handbook	0.015	0.033	0.019
	(0.013)	(0.029)	(0.030)
Assigned Handbook & Movie	0.030**	0.078***	0.055*
	(0.014)	(0.029)	(0.031)
Assigned Handbook & Counseling	0.026*	0.081***	0.070**
	(0.014)	(0.030)	(0.031)
Assigned All Three	0.016	0.082***	0.051*
	(0.014)	(0.030)	(0.030)
R-squared	0.113	0.142	0.127
N	2204	2205	2204
Dependent Variable Mean in Control Group	0.035	0.230	0.323
F-test (p-value): Book = Book & Movie	0.317	0.133	0.240
F-test (p-value): Book = Book & Counseling	0.442	0.120	0.107
F-test (p-value): Book = All Three	0.958	0.111	0.287
F-test (p-value): Book & Movie = Book & Counseling	0.813	0.920	0.645
F-test (p-value): Book & Movie = All Three	0.364	0.899	0.885

Notes: This table presents regression analysis for individual record keeping business practices. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), ** (5% significance level), and *** (1% significance level).

Appendix Table D3: Impact on Individual Planning Practices

	(1)	(2)	(3)
	Reviewed Financial Performance to Identify Areas of Improvement	Compared Target vs. Actual Monthly Sales	Anticipated Budget for Upcoming Business Costs
Assigned Handbook	0.009	0.034	0.037
	(0.033)	(0.035)	(0.024)
Assigned Handbook & Movie	0.038	0.053	0.036
	(0.033)	(0.035)	(0.023)
Assigned Handbook & Counseling	0.041	0.015	0.040*
	(0.032)	(0.033)	(0.023)
Assigned All Three	0.064**	0.081**	0.056**
	(0.032)	(0.034)	(0.023)
R-squared	0.108	0.130	0.185
N	2204	2204	2204
Dependent Variable Mean in Control Group	0.642	0.418	0.146
F-test (p-value): Book = Book & Movie	0.364	0.599	0.964
F-test (p-value): Book = Book & Counseling	0.311	0.571	0.914
F-test (p-value): Book = All Three	0.084	0.187	0.440
F-test (p-value): Book & Movie = Book & Counseling	0.942	0.263	0.876
F-test (p-value): Book & Movie = All Three	0.422	0.427	0.404

Notes: This table presents regression analysis for individual planning business practices. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), and ** (5% significance level).

Appendix Table D4: Impact on Individual Stocking Up Practices

	(1)	(2)	(3)	(4)
	Top Selling Products Always in Stock	Adjusted Stock Based on Product Profitability	Negotiated Lower Prices with a Supplier	Compared Product Prices and Quality Across Suppliers
Assigned Handbook	0.031	-0.000	-0.008	-0.044
	(0.029)	(0.033)	(0.028)	(0.034)
Assigned Handbook & Movie	0.065**	0.040	0.019	0.039
	(0.028)	(0.034)	(0.029)	(0.035)
Assigned Handbook & Counseling	0.026	-0.010	0.041	0.007
	(0.028)	(0.034)	(0.029)	(0.034)
Assigned All Three	0.041	0.091***	0.055*	0.038
	(0.029)	(0.034)	(0.029)	(0.034)
R-squared	0.055	0.111	0.108	0.155
N	2205	2205	2204	2204
Dependent Variable Mean in Control Group	0.770	0.420	0.195	0.500
F-test (p-value): Book = Book & Movie	0.212	0.223	0.343	0.016
F-test (p-value): Book = Book & Counseling	0.876	0.782	0.086	0.131
F-test (p-value): Book = All Three	0.696	0.007	0.026	0.014
F-test (p-value): Book & Movie = Book & Counseling	0.144	0.140	0.461	0.336
F-test (p-value): Book & Movie = All Three	0.385	0.135	0.222	0.965

Notes: This table presents regression analysis for individual stocking up business practices. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), ** (5% significance level), and *** (1% significance level).

Appendix Table D5: Impact on Individual Marketing Practices

	(1)	(2)
	Consulted with Former Customers	Offered Discount to Loyal/Bulk Customers
Assigned Handbook	0.008	-0.026
	(0.024)	(0.032)
Assigned Handbook & Movie	0.044*	0.043
	(0.026)	(0.034)
Assigned Handbook & Counseling	0.026	0.055
	(0.025)	(0.034)
Assigned All Three	0.056**	0.066*
	(0.027)	(0.034)
R-squared	0.099	0.145
N	2204	2205
Dependent Variable Mean in Control Group	0.142	0.358
F-test (p-value): Book = Book & Movie	0.181	0.037
F-test (p-value): Book = Book & Counseling	0.483	0.015
F-test (p-value): Book = All Three	0.077	0.005
F-test (p-value): Book & Movie = Book & Counseling	0.523	0.748
F-test (p-value): Book & Movie = All Three	0.677	0.508

Notes: This table presents regression analysis for individual marketing business practices. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), and ** (5% significance level).

Appendix Table D6: Impact on Individual Joint Decision Making Practices

	(1)	(2)
	Made Joint Business Decisions with Business Partner	Drafted an Agreement for Joint Decision-Making
Assigned Handbook	0.024	-0.001
	(0.031)	(0.027)
Assigned Handbook & Movie	0.038	0.041
	(0.031)	(0.028)
Assigned Handbook & Counseling	0.039	0.035
	(0.031)	(0.027)
Assigned All Three	0.068**	0.051*
	(0.030)	(0.028)
R-squared	0.126	0.107
N	2205	2205
Dependent Variable Mean in Control Group	0.310	0.228
F-test (p-value): Book = Book & Movie	0.642	0.128
F-test (p-value): Book = Book & Counseling	0.628	0.169
F-test (p-value): Book = All Three	0.149	0.057
F-test (p-value): Book & Movie = Book & Counseling	0.978	0.843
F-test (p-value): Book & Movie = All Three	0.322	0.719

Notes: This table presents regression analysis for individual joint decision making business practices. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), and ** (5% significance level).

Appendix Table D7: Quantile Regressions of Impact on Business Practices

	(1)
	Aggregate of All Practices
1st Decile:	
Assigned Handbook	-0.010
	(0.017)
Assigned Handbook & Movie	0.007
	(0.017)
Assigned Handbook & Counseling	0.014
	(0.015)
Assigned All Three	0.037**
	(0.018)
2nd Decile:	
Assigned Handbook	0.002
	(0.014)
Assigned Handbook & Movie	0.036**
	(0.016)
Assigned Handbook & Counseling	0.026**
	(0.013)
Assigned All Three	0.059***
	(0.015)
3rd Decile:	
Assigned Handbook	0.000
	(0.016)
Assigned Handbook & Movie	0.058***
	(0.015)
Assigned Handbook & Counseling	0.026**
	(0.013)
Assigned All Three	0.058***
	(0.014)
4th Decile:	
Assigned Handbook	0.026*
	(0.015)
Assigned Handbook & Movie	0.070***
	(0.014)
Assigned Handbook & Counseling	0.032**
	(0.013)
Assigned All Three	0.060***
	(0.013)
5th Decile:	
Assigned Handbook	0.015
	(0.014)
Assigned Handbook & Movie	0.053***
	(0.013)
Assigned Handbook & Counseling	0.020
	(0.014)
Assigned All Three	0.051***
	(0.014)
6th Decile:	
Assigned Handbook	0.013
	(0.015)
Assigned Handbook & Movie	0.057***
	(0.015)
Assigned Handbook & Counseling	0.026*
	(0.016)
Assigned All Three	0.057***
	(0.016)
7th Decile:	
Assigned Handbook	0.009
	(0.017)
Assigned Handbook & Movie	0.049***
	(0.016)
Assigned Handbook & Counseling	0.033*
	(0.017)
Assigned All Three	0.056***
	(0.017)
8th Decile:	
Assigned Handbook	0.000
	(0.020)
Assigned Handbook & Movie	0.032*
	(0.018)
Assigned Handbook & Counseling	0.039*
	(0.022)
Assigned All Three	0.052***
	(0.019)
9th Decile:	
Assigned Handbook	-0.002
	(0.022)
Assigned Handbook & Movie	0.033
	(0.021)
Assigned Handbook & Counseling	0.051**
	(0.022)
Assigned All Three	0.062***
	(0.023)
10th Decile	
Assigned Handbook	0.016
	(0.020)
Assigned Handbook & Movie	-0.011
	(0.020)
Assigned Handbook & Counseling	0.038*
	(0.022)
Assigned All Three	0.005
	(0.024)
N	2205

Notes: This table presents quantile regression analysis for the aggregate of all business practices. The output presents treatment effects for each decile. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. The specification includes a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are bootstrapped with 1,000 replications. Statistically significant p-values are highlighted by: * (10% significance level), ** (5% significance level), and *** (1% significance level).

Appendix E: Consistency of Estimated Profits Measure

This appendix explains the estimated profits measure used in this paper, and for robustness presents regression results using alternative measures.

In the main paper, we use an estimate of profits that is first derived as the difference between total sales and total expenses. Next, for values of profits that are negative, we replace them with the business owner's self-reported profits if the latter is available, positive, and the enumerator has expressed confidence in the business owner's answer.[2] We adopt this strategy based on the hypothesis that negative profits are likely to be unsustainable and if the two profit estimates contradict on the sign *and* the self-report is deemed reliable by enumerators, then we select the positive value. This data cleaning protocol was implemented at the time of the baseline survey and was carried through to ensuing data rounds.

We winsorize this value of profits on both tails at the 5% level (i.e. at the 5th and 95th percentiles) to arrive at the estimated profit measure used in the paper. Note that winsorizing the difference between sales and profits is slightly different than taking the difference between winsorized sales and winsorized expenses. The latter winsorizes the difference rather than levels. As before, this data cleaning protocol was adopted at the time of the baseline and followed through to later rounds.

Appendix Table E1 below presents the treatment impacts for all three measures of profits: (column 1) estimated profits with winsorized difference of sales and expenses (i.e. the measure we adopt); (column 2) estimated profits with the difference of winsorized values of sales and expenses; and (column 3) the difference of winsorized values of sales and expenses *without* the variable substitution described above.

The table shows that the treatment impacts for "Handbook & Counseling" and "All Three" are consistent and statistically significant across the three columns. The table also presents the autocorrelation coefficients and coefficients of variation (SD/Mean) for the control group across the three columns. The estimated profits measure in column (1) has the highest autocorrelation and lowest CV. For this reason and the fact that the estimated profits protocol was established at the time of the baseline survey, we use the column (1) estimated profits measure in the main paper.

Finally, Appendix Table E1 also presents the inverse hyperbolic sine transformation measure of estimated profits in column (4) as additional robustness. As with columns (2) and (3), the column (4) results are consistent with those in column (1).

[2] Our questionnaire included an enumerator assessment of the validity of self-reported answers to profits, as these estimates tend to be very noisy. For 52% of respondents, the enumerators expressed little or very little confidence in the business owners' answers. For this reason, we do not consider self-reported profits on its own as a reliable outcome measure.

Appendix Table E1: Alternative Measures of Profits

	(1)	(2)	(3)	(4)
	Estimated Profits Last Month (win 5%)	Sales (win 5%) - Expenses (win 5%)	Sales (win 5%) - Expenses (win 5%) *without* substituting negative values	Estimated Profits Last Month (IHS Transformation)
Assigned Handbook	-90.194	-165.269	-56.292	-0.067
	(78.342)	(110.148)	(140.481)	(0.088)
Assigned Handbook & Movie	113.842	78.746	77.886	0.055
	(86.829)	(123.485)	(150.669)	(0.092)
Assigned Handbook & Assistance	309.980***	295.979**	351.344**	0.261***
	(89.633)	(126.641)	(157.488)	(0.096)
Assigned All Three	190.267**	354.222**	455.552**	0.199**
	(84.814)	(158.305)	(181.539)	(0.094)
R-squared	0.180	0.122	0.064	0.211
N	2172	2172	2172	2172
Dependent Variable Mean in Control Group	894.544	1024.963	716.161	6.817
Dependent Variable SD in Control Group	1127.783	1668.285	2073.090	1.348
Dependent Variable Autocorrelation between Baseline and Endline in Control Group	0.37	0.32	0.09	0.33
Dependent Variable Coefficient of Variation (SD/Mean) in Control Group	1.26	1.63	2.89	0.20

Notes: This table presents regression analysis for alternative measures of business profits. Column (1) is the estimated profits measure used in the main paper, and the estimation procedure is described in Appendix E. Column (2) is winsorized sales minus winsorized expenses inclusive of the variable substitution described in Appendix E; and column (3) is winsorized sales minus winsorized expenses without the variable substitution. Finally, column (4) presents results for the inverse hyperbolic sine transformation measure of estimated profits. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), and ** (5% significance level).

Appendix F: Causal Mediation Analysis

Appendix Table F1 below presents results from formal causal mediation analysis to unpack the relative contribution of the business sub-practices. Moreover, it investigates the proportion of the treatment effects on profits, sales, and expenses in the *All Three* treatment arm that can be attributed to each dimension of practices. We acknowledge that this analysis requires a strong assumption on the independence of the mediator (i.e. business practices) from outcomes (i.e. business performance). Concretely, this assumption requires that there are no unobserved pre-treatment or *any* (observed or unobserved) post-treatment confounders that affect both the mediator and the outcome.

There are alternative methods for mediation analysis, such as sequential g-estimation (Acharya et al., 2016) that require less stringent assumptions. However, in order to recover the Average Direct Effect (ADE) and consequently the Average Causal Mediation Effect (ACME) we have to impose a different and non-trivial no-interactions condition (Carpena and Zia, 2020). There is also sensitivity analysis, as proposed by Imai, Keele, and Yamamoto (2010) and Imai, Keele, and Tingley (2010) that simulates the sensitivity of the ACME estimate to different degrees of violation of the identification assumption. Nevertheless, given the likelihood of unobservable confounders, this sensitivity analysis lacks an objective criterion for determining whether the identification assumption holds.

For these reasons, we do not emphasize or analyze the absolute magnitude of the aggregate practices index as a mediator. Instead, we find mediation analysis to be helpful when comparing the *relative* magnitudes of the different sub-practice channels. While still suggestive, this comparison relies on a slightly weaker identification assumption that there is no *differential* confounding effect across the different sub-practice mediators. Moreover, if we consider motivation or ability as an unobserved confounder, this assumption states that the confounder affects adoption of record keeping practices the same way it affects the adoption of other types of practices. As such, this comparison provides suggestive evidence on which business practice channels are more or less important.

Appendix F References:

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. “Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects.” *American Political Science Review*, 110:3, 512-529.

Carpena, Fenella, and Bilal Zia. 2020. “The Causal Mechanism of Financial Education: Evidence from Mediation Analysis.” *Journal of Economic Behavior & Organization*, 177:9, 143-184.

Imai, Kosuke, Luke Keele, and Dustin Tingley. 2010. “A General Approach to Causal Mediation Analysis.” *Psychological Methods*, 15:4, 309-334.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. “Identification, Inference and Sensitivity Analysis for Causal Mediation Effects.” *Statistical Science*, 25:1, 51-71.

Appendix Table F1: Causal Mediation Analysis

	(1)	(2)	(3)	(4)	(5)
	Record Keeping	Planning	Stocking-Up	Marketing	Joint Decision Making
Panel A: Profits					
ACME	7.257	7.805	3.863	15.667*	3.137
ADE	180.055**	179.599**	183.431**	171.827*	184.075**
Proportion of ATE Mediated	0.038	0.040	0.020	0.081	0.016
Panel B: Sales					
ACME	37.895	48.247	118.625**	187.702**	41.020
ADE	1119.352**	1109.821**	1040.302**	972.586**	1116.281**
Proportion of ATE Mediated	0.032	0.041	0.101	0.162	0.035
Panel C: Expenses					
ACME	46.760	47.139	137.650***	191.029***	47.565
ADE	741.186*	741.209*	651.928	599.688	740.331*
Proportion of ATE Mediated	0.057	0.058	0.169	0.234	0.058

Notes: This table presents results from formal causal mediation analysis for the All Three treatment group and sub-practice mediators. These mediators include indices of record keeping (column 1), planning (column 2), stocking up (column 3), marketing (column 4), and joint decision making (column 5). The table reports estimates for the average causal mediation effect (ACME), the average direct effect (ADE), and the proportion of the average treatment effect (ATE) that is mediated. Panel A presents estimates for profits; Panel B for sales, and Panel C for expenses. Statistically significant p-values are highlighted by: * (10% significance level), ** (5% significance level), and *** (1% significance level).

Appendix G: Gain/Loss Framing of the Handbook

Appendix Table G1 below presents treatment regressions on the main outcomes of this paper by the gain or loss framing of the handbook. As described in the main paper, half of the handbooks were distributed with economic returns to the adoption mentioned as gains and half as losses. At project inception, we were interested in differences in impact by gain and loss framing in the handbook where the former refers to practice adoption as untapped gains while the latter refers to lack of practice adoption as potential losses. The hypothesis was that loss aversion may lead to greater impacts when the handbook material is framed as losses rather than as gains.

While the results in Appendix Table G1 do show higher point estimates for the loss framing especially in practice adoption and for the “Handbook & Movie” treatment group, which is consistent with loss aversion, the coefficients for “Handbook & Counseling” as well as “All Three” are statistically indistinguishable across frames. For this reason and to maximize statistical power, we focus on pooled estimates in the main paper.

Appendix Table G1: Impact on Main Outcomes by Handbook Gain/Loss Framing

	(1)	(2)	(3)
	Aggregate of All Practices	Estimated Profits Last Month (win 5%)	Sales Last Month (win 5%)
Assigned Handbook (Gain Frame)	0.002	-4.666	-319.864
	(0.016)	(95.124)	(360.922)
Assigned Handbook (Loss Frame)	0.014	-171.432*	-442.724
	(0.015)	(94.574)	(384.315)
Assigned Handbook & Movie (Gain Frame)	0.022	-71.865	-22.106
	(0.015)	(99.122)	(419.836)
Assigned Handbook & Movie (Loss Frame)	0.063***	298.904**	755.511*
	(0.016)	(116.537)	(402.304)
Assigned Handbook & Counseling (Gain Frame)	0.024	348.192***	990.692**
	(0.016)	(116.295)	(497.274)
Assigned Handbook & Counseling (Loss Frame)	0.045***	272.326**	681.266
	(0.015)	(111.599)	(439.539)
Assigned All Three (Gain Frame)	0.043***	172.470*	640.577
	(0.016)	(104.398)	(443.598)
Assigned All Three (Loss Frame)	0.065***	206.769*	969.418**
	(0.016)	(108.297)	(449.051)
R-squared	0.311	0.185	0.495
N	2205	2172	2197
Dependent Variable Mean in Control Group	0.337	894.544	4998.923
Dependent Variable SD in Control Group	0.189	1127.783	5623.257
F-test (p-value): Book: Gain = Loss	0.523	0.121	0.763
F-test (p-value): Book & Movie: Gain = Loss	0.022	0.005	0.100
F-test (p-value): Book & Counseling: Gain = Loss	0.258	0.590	0.586
F-test (p-value): All Three: Gain = Loss	0.244	0.789	0.537

Notes: This table presents regression analysis for the aggregate of all practices, estimated profits, and sales by gain or loss framing of the handbook. Half the sample was randomly assigned a handbook with a gain frame and half with a loss frame. This table presents disaggregated analysis by frame type. The table also reports p-values from F-tests of equivalence between gain and loss frame within each treatment arm. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), ** (5% significance level), and *** (1% significance level).

Appendix H: Supplementary Tables

This appendix presents the following supplementary tables:

- Appendix Table H1: Impact on Business Practices by Facilitator/Beneficiary Gender Match
- Appendix Table H2: Impact on Unrelated Business Practices
- Appendix Table H3: Distance Analysis
- Appendix Table H4: Impact on Objective Measures
- Appendix Table H5: Impact on Main Outcomes by Education Level
- Appendix Table H6: Impact on Main Outcomes by Survey Round

Appendix Table H1: Impact on Business Practices by Facilitator/Beneficiary Gender Match

	(1)
	Aggregate of All Practices
Assigned Handbook	0.008
	(0.013)
Assigned Handbook & Movie	0.042***
	(0.013)
Book & Counseling: Gender Match	0.034**
	(0.016)
Book & Counseling: No Gender Match	0.035**
	(0.016)
All Three: Gender Match	0.049***
	(0.016)
All Three: No Gender Match	0.060***
	(0.017)
R-squared	0.308
N	2205
Dependent Variable Mean in Control Group	0.337
Dependent Variable SD in Control Group	0.189
F-test (p-value):Book & Counseling: Gender Match = No Gender Match	0.951
F-test (p-value): All Three: Gender Match = No Gender Match	0.589

Notes: This table presents regression analysis for business practices by facilitator and beneficiary gender match. The gender match = 1 if the counseling facilitator and beneficiary retailer have the same gender, and = 0 otherwise. The "Book & Counseling" and "All Three" treatment indicators are split by this gender matching. The table also reports p-values from F-tests of equivalence between gender match and no gender match. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: ** (5% significance level), and *** (1% significance level).

Appendix Table H2: Impact on Unrelated Business Practices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Aggregate of Unrelated Practices	Observed Products for Sale at Competing Business	Asked Suppliers about High Demand Products	Advertised the Business	Identified Profit Contribution of Best Products	Kept Annual Profit and Loss Statement	Kept Annual Cash Flow Statement	Kept Annual Balance Sheet	Kept Annual Income and Expenses Statement
Assigned Handbook	-0.003	0.007	-0.020	-0.003	0.005	0.007	-0.009	0.000	-0.010
	(0.010)	(0.031)	(0.032)	(0.013)	(0.030)	(0.011)	(0.011)	(0.008)	(0.013)
Assigned Handbook & Movie	-0.001	0.000	-0.035	0.012	0.020	0.002	0.001	-0.002	-0.004
	(0.010)	(0.031)	(0.033)	(0.015)	(0.029)	(0.012)	(0.012)	(0.008)	(0.013)
Assigned Handbook & Counseling	0.005	-0.000	-0.012	-0.005	0.014	0.013	0.013	0.005	0.013
	(0.010)	(0.031)	(0.032)	(0.013)	(0.029)	(0.011)	(0.013)	(0.009)	(0.014)
Assigned All Three	0.005	0.034	0.001	0.021	0.005	0.005	-0.008	-0.006	-0.001
	(0.009)	(0.031)	(0.032)	(0.014)	(0.029)	(0.012)	(0.010)	(0.007)	(0.013)
R-squared	0.183	0.107	0.137	0.088	0.080	0.075	0.067	0.036	0.103
N	2204	2204	2203	2204	2204	2204	2204	2204	2204
Dependent Variable Mean in Control Group	0.179	0.239	0.294	0.044	0.748	0.024	0.029	0.015	0.040
Dependent Variable SD in Control Group	0.132	0.427	0.456	0.206	0.435	0.154	0.167	0.124	0.196
F-test (p-value): Book = Book & Movie	0.814	0.839	0.626	0.331	0.633	0.679	0.347	0.760	0.612
F-test (p-value): Book = Book & Counseling	0.412	0.822	0.805	0.840	0.755	0.616	0.075	0.609	0.100
F-test (p-value): Book = All Three	0.365	0.414	0.508	0.095	0.995	0.867	0.905	0.377	0.476
F-test (p-value): Book & Movie = Book & Counseling	0.571	0.981	0.476	0.247	0.863	0.374	0.378	0.445	0.243
F-test (p-value): Book & Movie = All Three	0.519	0.305	0.265	0.546	0.625	0.807	0.389	0.596	0.825

Notes: This table presents regression analysis for business practices that were not included in the handbook. Column (1) present results for the aggregate measure, and columns (2)-(9) present results for individual practices. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level.

Appendix Table H3: Distance Analysis

	(1)	(2)
	Estimated Profits Last Month (win 5%)	Sales Last Month (win 5%)
Distance to nearest treatment shop (kms)	-476.227	-1445.029
	(593.468)	(2579.911)
Distance to nearest control shop (kms)	98.361	-1001.081
	(287.183)	(1233.178)
R-squared	0.248	0.530
Observations	445	452

Notes: This table investigates spillover effects among the control group by regressing performance outcomes of profits (column 1) and sales (column 2) on the distance to the nearest treatment shop. To account for market density, the regressions also control for distance to the nearest control shop. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level.

Appendix Table H4: Impact on Objective Measures

	(1)	(2)	(3)	(4)	(5)
	Objective Measures Aggregate Score	Shop is Clean and Tidy	Shop is Well-Stocked	Prices are Clearly Marked	Goods are Displayed in Groups
Assigned Handbook	0.005	-0.016	0.017	0.037	-0.009
	(0.015)	(0.026)	(0.030)	(0.025)	(0.016)
Assigned Handbook & Movie	0.017	-0.029	0.070**	0.052**	-0.015
	(0.016)	(0.027)	(0.029)	(0.025)	(0.018)
Assigned Handbook & Counseling	0.004	-0.031	0.029	0.043*	-0.011
	(0.015)	(0.027)	(0.030)	(0.025)	(0.017)
Assigned All Three	0.013	-0.015	0.024	0.034	0.014
	(0.015)	(0.027)	(0.031)	(0.024)	(0.016)
R-squared	0.114	0.148	0.119	0.065	0.063
N	2204	2204	2204	2204	2204
Dependent Variable Mean in Control Group	0.609	0.673	0.695	0.144	0.925
Dependent Variable SD in Control Group	0.237	0.470	0.461	0.351	0.264
F-test (p-value): Book = Book & Movie	0.449	0.615	0.072	0.570	0.748
F-test (p-value): Book = Book & Counseling	0.976	0.571	0.685	0.830	0.918
F-test (p-value): Book = All Three	0.596	0.977	0.827	0.890	0.159
F-test (p-value): Book & Movie = Book & Counseling	0.433	0.957	0.171	0.722	0.827
F-test (p-value): Book & Movie = All Three	0.795	0.609	0.119	0.472	0.105

Notes: This table presents regression analysis for objective measures that enumerators noted during survey interviews. Column (1) present results for the aggregate measure, and columns (2)-(5) present results for individual measures. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), and ** (5% significance level).

Appendix Table H5: Impact on Main Outcomes by Education Level

	(1)	(2)	(3)	(4)	(5)	(6)
	Aggregate of All Practices			Estimated Profits Last Month (win 5%)		
	Primary or Less	Primary to High School	More than High School	Primary or Less	Primary to High School	More than High School
Assigned Handbook	0.004	0.030	0.002	-8.922	-84.091	-71.383
	(0.022)	(0.029)	(0.019)	(138.770)	(193.273)	(115.116)
Assigned Handbook & Movie	0.082***	0.062**	0.014	157.993	76.261	120.770
	(0.021)	(0.029)	(0.020)	(146.439)	(233.299)	(128.758)
Assigned Handbook & Counseling	0.032	0.052*	0.026	462.736***	302.719*	313.715**
	(0.023)	(0.027)	(0.019)	(168.244)	(175.741)	(129.873)
Assigned All Three	0.089***	0.039	0.036*	313.489*	82.880	216.000*
	(0.021)	(0.030)	(0.019)	(162.506)	(201.462)	(121.658)
R-squared	0.326	0.319	0.327	0.198	0.286	0.196
N	661	468	1076	653	462	1057
Dependent Variable Mean in Control Group	0.281	0.325	0.388	840.559	939.636	917.854
Dependent Variable SD in Control Group	0.167	0.175	0.198	998.037	1021.182	1266.246
F-test (p-value): Book = Book & Movie	0.001	0.312	0.513	0.294	0.513	0.114
F-test (p-value): Book = Book & Counseling	0.252	0.388	0.208	0.007	0.049	0.003
F-test (p-value): Book = All Three	0.000	0.747	0.072	0.068	0.436	0.016
F-test (p-value): Book & Movie = Book & Counseling	0.041	0.727	0.556	0.094	0.309	0.170
F-test (p-value): Book & Movie = All Three	0.789	0.470	0.251	0.380	0.979	0.449

Notes: This table presents regression analysis for the aggregate of all practices and estimated profits, separately by baseline education levels: primary or less, primary to high school, and greater than high school. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), ** (5% significance level), and *** (1% significance level).

Appendix Table H6: Impact on Main Outcomes by Survey Round

	(1)	(2)	(3)	(4)	(5)	(6)
	Aggregate of All Practices		Estimated Profits Last Month (win 5%)		Sales Last Month (win 5%)	
	Midline	**Endline**	**Midline**	**Endline**	**Midline**	**Endline**
Assigned Handbook	0.007	0.009	-116.411	-67.301	-355.311	-400.731
	(0.015)	(0.016)	(90.399)	(107.204)	(338.271)	(382.431)
Assigned Handbook & Movie	0.055***	0.026*	76.966	157.616	661.948*	28.309
	(0.016)	(0.016)	(97.121)	(120.141)	(367.536)	(411.670)
Assigned Handbook & Counseling	0.039***	0.029*	273.112***	353.037***	942.757**	723.637
	(0.015)	(0.016)	(98.753)	(123.669)	(391.743)	(453.300)
Assigned All Three	0.058***	0.049***	141.481	254.038**	696.922*	931.730**
	(0.015)	(0.016)	(97.913)	(118.024)	(386.365)	(431.161)
R-squared	0.322	0.362	0.209	0.172	0.534	0.462
N	1181	1024	1177	995	1179	1018
Dependent Variable Mean in Control Group	0.327	0.348	896.854	891.815	5024.003	4970.021
Dependent Variable SD in Control Group	0.184	0.193	1106.399	1155.260	5572.596	5694.277
F-test (p-value): Book = Book & Movie	0.002	0.272	0.044	0.056	0.004	0.245
F-test (p-value): Book = Book & Counseling	0.031	0.220	0.000	0.001	0.000	0.006
F-test (p-value): Book = All Three	0.001	0.015	0.008	0.006	0.004	0.001
F-test (p-value): Book & Movie = Book & Counseling	0.302	0.877	0.057	0.145	0.480	0.112
F-test (p-value): Book & Movie = All Three	0.887	0.173	0.529	0.453	0.929	0.031

Notes: This table presents regression analysis for the main outcomes of this paper, separately by survey round. Columns (1) and (2) present results for the aggregate practices measure; columns (3) and (4) for estimated profits; and columns (5) and (6) for sales. Odd numbered columns present analysis from the midline, and even numbered columns present analysis from the endline. All regression specifications control for the baseline value of the outcome variable, as well as stratification controls. Robust standard errors are reported in parentheses. Statistically significant p-values are highlighted by: *(10% significance level), ** (5% significance level), and *** (1% significance level).

Appendix I: Handbook Business Practices

This appendix presents the following descriptive tables:

- Appendix Table I1: List of Business Practices Measured in Surveys
- Appendix Table I2: Business Practices Included in the Handbook

Appendix Table I1: List of Business Practices Measured in Surveys

Label	Name of the Practice	Survey Question
Record Keeping:		
R1	Keeps written business records	Do you keep written business records?
R2	Records every purchase and sale	Do you record each purchase and sale of your business?
R3	Can use records to check cash on hand	Are you able to use your records to see how much cash your business has available at any point in time?
R4	Uses records to check sales of product	Do you regularly use your records to know whether sales of a particular product are increasing or decreasing?
R5	Works out costs of main products	Do you work out the costs to the business of each main product you sell?
R6	Knows products with most profit per item	Do you know your three products with the highest profit per item selling in your business?
R7	Has written monthly expenses budget	Do you have a written budget which states how much you have to pay each month for rent, electricity, equipment maintenance, transport, advertising, and other indirect costs to business?
R8	Can use records to service potential loan	Imagine you wanted to take out a bank loan for your business and you needed to check if you can pay the monthly rates. Do you have records showing whether you have enough money each month after paying all business expenses?
R9	Keeps records in ledger book	What kind of business records do you keep? a) Electronic records (phone, computer, etc.) b) Ledger book or equivalently organized notes c) Disorganized notes d) Orderly filed receipts e) Loose receipts f) Other g) None
R10	Calculates profits considering all costs	How do you calculate profits? a) Only from the sales (no calculation) b) Sales minus product purchases c) Sales minus product purchases minus some other business expenses d) Sales minus all business expenses (salaries, rent, electricity, interest, etc.)
Planning:		
F1	Reviews and analyses financial performance	Do you review your financial performance and analyze where there are areas for improvement at least on a monthly basis?
F2	Has sales target set for next year	Do you have a target set for sales over the next year?
F3	Compares target with sales at least monthly	Do you compare your sales achieved to your target at least monthly?
F4	Has cost budget for next year	Do you have a budget for the anticipated costs your business will have to pay over the next year?
F5	Issues annual profit/loss statement	Do you have an annual profit and loss statement?
F6	Issues annual cash-flow statement	Do you have an annual statement of cash flow?
F7	Issues annual balance sheet	Do you have an annual balance sheet?
F8	Issues annual income and expenditure sheet	Do you have an annual income/expenditure statement?

Appendix Table I1: List of Business Practices Measured in Surveys

Label	Name of the Practice	Survey Question
Stocking Up:		
S1	Negotiates prices with suppliers	In the last 3 months, have you tried to negotiate with any of your suppliers for a lower price of products?
S2	Compares suppliers on product quality	In the last 3 months, have you compared the prices or quality offered by alternate suppliers to the business' current suppliers?
S3	Does not run out of stock	Does your business run out of stock monthly or more often?
S4	Stocks up late for any of top 3 selling products	How often do you stock up your top 3 products? a) No fixed schedule, whenever products are out of stock b) No fixed schedule, whenever products are almost out of c) Fixed schedule, plus when products are out of stock d) Fixed schedule, plus when products are almost out of stock e) Whenever an item or product is sold
S5	Changes inventory based on profits per item	Do you ever adjust your level of inventory based any of the following reasons? a) Change in customer demand b) Shelf space c) Purchasing prices d) Profits per item
S6	Changes inventory based on buying price	(see above: *S5*)
Marketing:		
M1	Visits competitors to see prices	In the last 3 months, have you visited at least one of your competitors' businesses to see their selling price?
M2	Visits competitors to see products	In the last 3 months, have you visited at least one of your competitors' businesses to see what products they have available for sale?
M3	Asks customers for new products they would like	In the last 3 months, have you asked existing customers if there are other products or brands they would like you to sell?
M4	Asks former customers why they quit buying	In the last 3 months, have you talked with at least one former customer to find out why they have stopped buying from your shop?
M5	Asks suppliers for bestselling products	In the last 3 months, have you asked any of your supplier about which products sell the best?
M6	Attracts customers with special offer	In the last 3 months, have you tried to attract customers with a special offer?
M7	Advertises the business	In the last 6 months, have you ever advertised your shop in any way?
M8	Compares sales with competitors	Do you compare your own firm's sales performance with your competitors?
M9	Adjusts prices based on competitors	In the last 3 months, have you changed prices for any of the following reasons? a) Different price charged by competitor b) More/less buyers than normal for that product c) Discount for bulk purchases d) Discount for loyal customers e) Discount for stocks in need to be sold quickly
M10	Gives any discount	(see above: *M9*)
M11	Has discussions about new products	What kind of business topics do you discuss with other people? a) Sales b) Selling price c) Bestselling products d) Discounts e) Promotion, marketing, advertising f) Government funding g) Other financing opportunities h) ROSCAs i) Supplier(s) j) Purchasing prices k) New brands or products l) Business practices m) Business plan n) Market trends o) Business rumors p) Security-related issues q) Most profitable products r) Assets owned s) New assets t) Others
M12	Has discussions about suppliers	(see above: *M11*)
M13	Has discussions about bestsellers	(see above: *M11*)

Appendix Table I1: List of Business Practices Measured in Surveys

Label	Name of the Practice	Survey Question
Joint Decision-Making:		
J1	Discusses business with family members	With whom do you discuss your business most often? a) Family member b) Personal friend c) Business friend in neighborhood d) Business friend outside of neighborhood e) Supplier f) Local official ("Kepala RT/RW/Kelurahan") g) Preman h) No one
J2	Discusses business with business friends	(see above: *J1*)
J3	Discusses business matters with anyone	(see above: *J1*)
J4	Discusses business sales	What kind of business topics do you discuss with other people? a) Sales b) Selling price c) Best-selling products d) Discounts e) Promotion, marketing, advertising f) Government funding g) Other financing opportunities h) Arisan ("ROSCAs") i) Supplier(s) j) Purchasing prices k) New brands or products l) Business practices m) Business plan n) Market trends o) Business rumors p) Security-related issues q) Most profitable products r) Assets owned s) New assets t) Others
J5	Discusses selling prices	(see above: *J4*)
J6	Discusses best-selling products	(see above: *J4*)
J7	Discusses financing opportunities	(see above: *J4*)
J8	Discusses buying prices	(see above: *J4*)
J9	Discusses business practices	(see above: *J4*)
J10	Discusses business plan	(see above: *J4*)
J11	Joint decisions on business sales	What kind of business decisions do you make with other people? a) Sales b) Selling price c) Best-selling products d) Discounts e) Promotion, marketing, advertising f) Government funding g) Other financing opportunities h) Arisan ("ROSCAs") i) Supplier(s) j) Purchasing prices k) New brands or products l) Business practices m) Business plan n) Market trends o) Business rumors p) Security-related issues q) Most profitable products r) Assets owned s) New assets t) Others
J12	Joint decisions on selling prices	(see above: *J11*)
J13	Joint decisions on best-sellers	(see above: *J11*)
J14	Joint decisions on financing opportunities	(see above: *J11*)
J15	Joint decisions on buying prices	(see above: *J11*)
J16	Joint decisions on new products	(see above: *J11*)
J17	Joint decisions on business practices	(see above: *J11*)
J18	Joint decisions on business plan	(see above: *J11*)

Notes: This table lists the business practices measured in this study, along with the corresponding survey question for each practice.

Table I2: Business Practices Included in the Handbook

	Number of Significant Specifications at Baseline	Baseline OLS Coefficients	
		Monthly Sales	Monthly Profits
	(1)	(2)	(3)
Marketing			
Asks former customers why they quit buying	7	0.245	0.226
Gives any discount	8	0.401	0.294
Stocking up			
Stocks up late for any of top 3 products	8	-0.250	-0.260
Has fixed stock-up schedule for top 3 products	6	0.263	0.247
Stocks up top 3 at most weekly	8	-0.483	-0.366
Record-keeping			
Keeps written business records	7	0.283	0.259
Can use records to service potential loan	8	0.351	0.358
Records outstanding payments by customers	7	0.401	0.361
Calculates profits considering all costs	8	0.418	0.445
Financial Planning			
Introduced at least 5 new products	8	0.354	0.279
Joint Decision-making			
Discusses business matters with anyone	8	0.253	0.272
Joint decisions on new products	8	0.299	0.345
Joint decisions on business practices	8	0.256	0.270

Note: This table lists the business practices that were included in the handbook. Column (1) reports the number of baseline outcome specifications where the practice was significantly correlated with the outcome. Columns (2) and (3) report the corresponding coefficients for regressions with sales and profits as dependent variables, respectively.

Appendix J: Comparison with Other Studies

This appendix presents the following descriptive table:

- Appendix Table J1: Comparison with Other Studies

Appendix Table J1: Comparison with Other Studies

Study	Mentoring	Training	Consulting	Other	Profits		Sales		Longest Endline Survey
	(Length in Hours or Days)				*%*	*St. Dev.*	*%*	*St. Dev.*	*(months)*
Field, Jayachandran, & Pande (2010)		2 days			n.r.	n.r.	n.r.	n.r.	4
Karlan & Valdivia (2011)		8.5 or 22			no effect	n.r.	no effect	n.r.	12 to 24
Bruhn & Zia (2011)		6*			no effect	n.r.	n.r.	n.r.	12
Klinger & Schundeln (2011)		7 days			n.r.	n.r.	n.r.	n.r.	12
Sonobe, Suzuki, & Otsuka (2011)									
Tanzania		50[1]			no effect	n.r.	no effect	n.r.	4
Ethiopia		37.5[2]			n.r.	n.r.	n.r.	n.r.	4
Vietnam-Steel		32.5[3]			n.r.	n.r.	n.r.	n.r.	4
Vietnam-Knitwear		32.5[3]			n.r.	n.r.	n.r.	n.r.	4
Berge, Bjovart & Tungodden (2012)									
Male		15.75			no effect	n.r.	**31.0%**	n.r.	29 to 31
Female		15.75			no effect	n.r.	no effect	n.r.	29 to 31
De Mel, McKenzie, & Woodruff (2012)									
Current Enterprises		49 to 63*			no effect	n.r.	no effect	n.r.	16 to 25
Potential Enterprises		49 to 63*			**43.1%**	n.r.	**40.9%**	n.r.	16 to 25
Mano, Iddrisu, Yoshino, & Sonobe (2012)		37.5			no effect	n.r.	no effect	n.r.	12
Valdivia (2015)									
General Training		108			n.r.	n.r.	no effect	n.r.	19-25
General Training and Technical Assistance		108			n.r.	n.r.	**20.4%**	n.r.	19-25
Premand, Brodmann, Almeida, Grun, & Barouni (2016)		20 days +			n.r.	n.r.	n.r.	n.r.	9 to 12
Calderon, Cunha, & De Giorgi (2013)		48			**24.0%**	n.r.	**28.5%**	n.r.	8 and 28
Bloom, Eifert, Mahajan, McKenzie and Roberts (2013)			508[4]		n.r.	n.r.	n.r.	n.r.	36
Giné & Mansuri (2014)		46			no effect	n.r.	no effect	n.r.	19-22
Glaub, Frese, Fischer, & Hoppe (2014)		3 days			n.r.	n.r.	**44.7%**	n.r.	12
Drexler, Fischer & Schoar (2014)		15 to 18			n.r.	n.r.	no effect	no effect	12
Fairlie, Karlan, & Zinman (2015)		8.5[5]			no effect	n.r.	n.r.	n.r.	60
Sayinzoga, Bulte, & Lensink (2015)		45[6]			n.r.	n.r.	n.r.	n.r.	15
Karlan, Knight & Udry (2015)			10		n.r.	n.r.	no effect	no effect	12
Bulte, Lensik & Vu (2017)		9[7]		9-18[4]	no effect	no effect	no effect	no effect	12
Lafortune, Riutort, & Tessada (2018)									
Role Model		48 to 56[8]			**31.1%**	**0.17**	**16.7%**	**0.10**	12
Technical Assistance (Group)		48 to 56[8]			no effect	no effect	no effect	no effect	12
Technical Assistance (Individual)		48 to 56[8]			no effect	no effect	**33.4%**	**0.19**	12
Cai & Szeidl (2018)				144	**24.7%**	n.r	**8.1%**	n.r.	12 to 24
Anderson, Chandy & Zia (2018)		80			**61.0%**	**0.3**	**64.0%**	**0.3**	12
Bruhn, Karlan & Schoar (2018)			208[9]		no effect	no effect	no effect	no effect	15 to 18
Brooks, Donovan, & Johnson (2018)									
Business Class		8[10]			no effect	n.r.	no effect	n.r.	17
Mentorship					**20.6%**	n.r.	no effect	n.r.	17
This paper				1	**35%**	**0.28**	**16.7%**	**0.15**	18

Notes: This table summarizes the study design and business performance impacts across the business training/mentorship/consulting literature. The table reports the duration of each intervention, the impact on profits and sales, and the length of the measurement period. The table reports impacts of the most effective treatment in each study.

Other Notes: n.r. denotes not reported. (1) 20 working days times 2.5 hours per day (class training). (2) 15 working days times 2.5 hours per day (class training). (3) 13 working days times 2.5 hours per day (class training). (4) 781 hours for treatment less 273 hours for control. (5) Difference in additional hours of training for treatment compared to control. (6) 5 days at 9 hours. (7) 1 hour module for 9 months and 36 weeks of meetings at 15-30 minutes. (9) Consultants provided services for 52 weeks at 4 hours per week. (10) 4 classes 2 hours each. n.r. not reported. Finally, a * indicates that the figure was taken from McKenzie and Woodruf (2013).

References for Appendix Table J1

Anderson, S. J., Chandy, R., & Zia, B. (2018). Pathways to profits: The impact of marketing vs. finance skills on business performance. *Management Science* , 64(12), 5559-5583.

Berge, L., Bjorvatn, K., Juniwaty, K., & Tungodden, B. (2012). Business training in Tanzania: From research driven experiment to local implementation. *Journal of African Economies* , 21(5): 808-827.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does Management Matter? Evidence from India. *Quarterly Journal of Economic* , 128(1):1-51.

Brooks, W., Donovan, K., & Johnson, T. (2018). Mentors or Teachers? Microenterprise Training in Kenya. *American Economic Journal: Applied Economics* , 10(4):196-221.

Bruhn, M., & Zia, B. (2011). Stimulating managerial capital in emerging markets: the impact of business and financial literacy for young entrepreneurs. *Journal of Development Effectiveness* , 5(2): 232-266.

Bruhn, M., Karlan, D., & Schoar, A. (2018). The Impact of Consulting Services on Small. *Journal of Political Economy* , 126(2): 635-687.

Bulte, E., Lensink, R., & Vu, N. (2017). Do gender and business trainings affect business outcomes? Experimental evidence from Vietnam. *Management Science* , 63(9), 2885-2902.

Cai, J., & Szeidl, A. (2018). Interfirm relationships and business performance. *The Quarterly Journal of Economics* , 133(3), 1229-1282.

Calderón, G., Cunha, J., & De Giorgi, G. (2013). Business Literacy and Development: Evidence from a Randomized Controlled Trial in Rural Mexico. Barcelona GSE Working Paper Series, 742.

De Mel, S., McKenzie, D., & Woodruff, C. (2012). Business Training and Female Enterprise Start-up, Growth, and Dynamics: Experimental Evidence from Sri Lanka. World Bank.

Drexler, A., Fischer, G., & Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics* , 6(2), 1-31.

Fairlie, R., Karlan, D., & Zinman, J. (2015). Behind the gate experiment: Evidence on effects of and rationales for subsidized entrepreneurship training. *American Economic Journal: Economic Policy,* 7(2): 125-161.

Field, E., Jayachandran, S., & Pande, R. (2010). Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India. *American Economic Review Papers and Proceedings* , 100 (2): 125–29.

Giné, X., & Mansuri, G. (2014). Money or Management? A Field Experiment on Constraints to Entrepreneurship in Rural Pakistan. World Bank Working Paper Series, 6959.

References for Appendix Table J1

Glaub, M., Frese, M., Fischer, S., & Hoppe, M. (2014). Increasing Personal Initiative in Small Business Managers or Owners Leads to Entrepreneurial Success: A Theory-Based Controlled Randomized Field Intervention for Evidence-Based Management. *Academy of Management Learning & Education* .

Karlan, D., & Valdivia, M. (2011). Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions. *Review of Economics and Statistics* , 93 (2): 510–27.

Karlan, D., Knight, R., & Udry, C. (2015). Consulting and capital experiments with microenterprisetailors in Ghana. *Journal of Economic Behavior & Organization* , 281-302.

Klinger, B., & Schündeln, M. (2011). Can Entrepreneurial Activity be Taught? Quasi-Experimental Evidence from Central America. World Development, 39 (9): 1592–610.

Lafortune, J., Riutort, J., & Tessada, J. (2018). Role Models or Individual Consulting: The Impact of Personalizing Micro-entrepreneurship Training. American Economic Journal: Applied Economics,10(4):222-245.

Mano, Y., Iddrisu, A., Yoshino, Y., & Sonobe, T. (2012). How Can Micro and Small Enterprises in Sub-Saharan Africa Become More Productive? The Impacts of Experimental Basic Managerial Training. *World Development* , 40(3): 458–68.

Premand, P., Brodmann, S., Almeida, R., Grun, R., & Barouni, M. (2016). Entrepreneurship education and entry into self-employment among university graduates. *World Development* , 77, 311-327.

Sayinzoga, A., Bulte, E., & Lensink, R. (2014). Financial Literacy and Financial Behaviour: Experimental Evidence From Rural Rwanda. *The Economic Journal* , 126(594): 1571-1599.

Sonobe, T., Suzuki, A., & Otsuka, K. (2011). Kaizen for Managerial Skills Improvement in Small and Medium Enterprises: An Impact Evaluation Study. World Bank.

Valdivia, M. (2015). Business training plus for female entrepreneurship? Short and medium-term experimental evidence from Peru. *Journal of Development Economics* , 113, 33-51.

Appendix K: Concept and Content of the Handbook

This appendix presents summary concepts in each chapter of the handbook.

Appendix K: Handbook Content on Beliefs, Reasons to Adopt Practices, Step-by-Step Implementation Guidance, and Tips

Chapter 1: Record Keeping

Misconception	HB Response
Some people think keeping records is difficult for people without higher education.	In our survey, those who keep records of their daily transactions varied widely in their educational background.
Some people think keeping records is complicated.	We will show you easy steps to take in order to keep track of your daily transactions
Many people believe they know their profits, know how much to reinvest, and how much to spend.	In our research we learned that many shop owners did not have a good idea of their business performance.
Some people think that one cannot prepare for unexpected events since nobody can predict the future.	Record keeping is a good example of how you can, in fact, prepare yourself. It makes sense to save up for large bills which arrive periodically, for instance, electricity and telephone bills, as well as rental costs and taxes.
Some people think keeping records is a lot of work and yields no benefits.	We will show you that once you know how to do it and how to integrate it into your everyday routines, record-keeping is a lot easier than it might first seem. In our survey, we found that a majority of people confuse business sales with profits. As a consequence, they may end up spending more than they actually have and thus risk running into financial trouble.

Reasons to Adopt the Practice
It is what differentiates successful businesses from those with less success.
Once you do record-keeping, you will know exactly how much you make per day, how much you can reinvest in your shop, how much you can save, and how much you can spend on you and your family.
Without savings, unexpected events can hit you and your family hard.
It allows you to take full control over all business decisions.
It helps keeping household and business finances separate.
You will be better able to save up and withstand unexpected events.
It is essential when borrowing from banks, money lenders, and family. Often, the only way to convince a potential lender is by showing them a written record of how your business is performing. Keeping records of your daily sales and weekly or monthly profits does just that. It gives you proof of how well your shop is performing and that the money will be paid back in time.

Chapter 1: Record Keeping

Step-by-step Implementation Guidance
Step I: Physically separating private and business finances (use labels)
Step 2: Preparing the ledger book (Column I: “Product", Column 2: "Amount"…)
Step 3: Recording daily sales (with tips: e.g. leave the first row on the first page of your book empty; with detailed examples on how to do it)
Step 4: Recording stock-up purchases.
Step 5: Dealing with household expenses (with concrete realistic examples: "your kid takes a snack or your aunt borrow money" and local tips "use colored pen for household expenses")
Step 6: Dealing with stalled payments (“paying later“) with concrete local example "let’s say when Pak Agung (a loyal customer) comes to your shop to buy 10 cans of condensed milk, he does not have enough money to complete the purchase."
Step 7: Adding other costs (electricity, transportation, rent, etc.) (tips: used a colored pen, divide the monthly value per days in the month, keep the invoice either behind the current page or at the end of the ledger.)
Step 8: Summing up at the end of the page (tips: leave two rows free at the end of each page. Draw a line beneath the last entry and use these last rows to note down the total sum of money received in column four (“Cash in”) and the total sum of money spent in column five (“Cash out”).
Step 9: Balance at the end of the day (and for the next).

Tips
Dealing with sales you forgot to write down: Whenever you realize you have forgotten a transaction, leave a few lines empty and immediately fill in the information you remember. For instance, you may remember the product name and unit (e.g. “Shampoo, 25ml sachets”). Should you fail to remember anything, simply leave one single line empty and carry on.
Dealing with electricity bills (vouchers): Switch to a voucher-based payment system, which allows you to top up your electricity balance when you would like to with exactly the amount you have allocated for it.
Set aside money to pay regular bills (electricity, rent, etc.).
Dedicate a third wallet/drawer/box to savings for these long-term expenses and to label it clearly (for instance, “savings for electricity bill”).

Chapter 2: Calculating Profits

Misconception	HB Response
Some people think that counting sales is already enough.	When we did our survey back in March 2016, we asked detailed questions about how people dealt with costs, sales, profits and record-keeping in general. We found that many people gave up calculating profits and some just consider sales as a guide for the performance of their business. The HB shows by means of an idiosyncratic example the consequences of this mistake.
Some people think that profits are simply “sales minus product purchases”.	The HB shows by means of an idiosyncratic example the consequences of this mistake.
Some people think that profits are “sales minus everything you pay that day”.	The HB shows by means of an idiosyncratic example the consequences of this mistake.
Some people think calculating profits is complicated.	The HB shows it can be simple.

Reasons to Adopt the Practice
Shops which have records detailed enough that they allow for the shop manager to calculate profits including all major costs to the business are associated with monthly sales 42% higher and profits 45% higher than shops without detailed records or without any records at all.
Only if you calculate profits, you can draw the right conclusions from your records.
You may overestimate your profits as you do not account for various bills and expenses. As a consequence, you may spend more money than you should, especially if there are outstanding bills to be paid in the future. Conversely, you may have the impression you are short on money, even if that is not the case. The HB gives an idiosyncratic example to illustrate both cases.

Step-by-step Implementation Guidance
Step 1: Calculating total daily sales and costs at the end of the day. The HB goes back to Step 9 of record-keeping. It uses idiosyncratic rules of thumb like "the sum where you disregard all colored entries will now be used to calculate profits. Entries highlighted with color represent costs that are not related to business expenses, such as loans to your sister". The HB illustrates with the example carried from Chapter 1
Step 2: Calculating daily profits "simply take your total daily sales and subtract from it your total daily costs." Illustrates with the example carried from Chapter 1
Step 3: Calculating monthly profits and profits on a “normal day”. "calculating your monthly profits is equivalent to calculating profits on a daily basis. Hence, all you have to do is add up your daily profits for all days of the month."

Tips
Making daily profits more precise (stock-ups for multiple days): "With the records now in place, you can estimate the days it takes until the products stocked up on one occasion sell out completely....Try to estimate the number of days for each product to stock-out. The way you deal with the entries in your records will be the same as with costs that represent more than one day. Just as you have done already (record-keeping, step 7), every purchase that represents more than a day shall be marked with color. Then add the total purchases for those products divided by the days they will last for."
Matching daily profits with monthly profits: "at the end of a month, calculate the monthly profits as shown above. Then, go back to each day of the month and note down, on the page that you are calculating your monthly profits, both the total daily sales and that day’s total costs including those you marked with color. If you now sum up the daily total sales, you receive your monthly total sales. Likewise, you can calculate the monthly total costs."

Chapter 3: Stocking Up

Misconception	HB Response
In our survey, we found that concerning stock-ups some people go by the principle that stocking up is not necessary until things are lacking on your shelf.	You lose money with every customer who decides to turn to another shop because they cannot wait for you to stock up. If you want to retain customers, it is vital to schedule your product purchases.
"I simply stock up the exact same amount that I had on my shelves before"	If you do this, you can run out of stock on special days (like Lebaran and other holidays).
"I have my favourite supplier with the best prices and do not need any other"	A shop owner who relies entirely on one or two suppliers, no matter how good the prices, will become dependent. When we asked owners of small and mid-sized stores to share advice on doing business, they stressed two things: not to rely on one or two suppliers only and to keep up to date with prices.

Reasons to Adopt the Practice
Stores whose owners kept a schedule to plan when to stock up and how much were associated with 26% higher monthly sales and 25% higher monthly profits.
However, most businesses do not plan ahead for supplies. Hence, they will be more prone to experience stock-outs. Shops which stock up daily are associated with monthly sales that are 48% higher and monthly profits that are 37% higher.
With a ready plan for stock-ups you will not have to purchase supplies every day and you will still be able to adjust when things change. We found that stores which never ran out of stock for their main products are associated with 25% higher monthly sales and 27% higher monthly profits.
You will be better able to pick up on product trends or seasonal changes for a particular product. Likewise, you will be able to observe the sales of new products more closely.
A fixed schedule will give you a better focus in managing your business and certainly make things easier for you. You do not make yourself dependent on your supplier's schedule, but follow your own. If your stock-up schedule is not fixed, you will tend to stock up more according to what is available at your suppliers. With your own schedule you will know better when to purchase any given product.

Chapter 3: Stocking Up

Step-by-step Implementation Guidance
Step 1: Preparing the material. It is best to keep your records in a special book (your inventory book). Any booklet or notebook will do.
Step 2: Figure out your most profitable products. Select the five to seven most profitable products, since it will be easier to track every single product's stock-out (step 3), demand (step 4), compare its requirements for storage, capital, and transport relative to other products (step 5), and compare relevant prices, product quality, availability, and supplier schedules (step 6).
Step 3: Get an idea about stock-outs. Take the ledger where you listed your most profitable products in (step 1) and start taking note of how often you run out of stock for these products. If you ask your customers how much exactly they were going to buy, you will have to note this down in a certain unit. For example, if a customer asks for 4 bottles of soy sauce, you can add 4 tallies in the row of that product in your records. Label the third column with "Unit" and the fourth column with "Stock-outs". In the fourth column you will keep the tallies for stockouts of all your main products. The examples in the HB illustrate all this with typical goods sold.
Step 4: Get an idea of your inventory. Add the following two new column labels: label column 5 with "Items sold" and column 6 with "Inventory". In column 5, you will keep count of the number of items sold for each of the products using a tally in just the same way you did with stock-outs.
Step 5: Identify constraints (storage, capital, transport, limited shelf-life, manpower). The HB proposes product-specific solutions for typical problems with stocking-up (e.g. lack of storage space, size of vehicle, and time)
Step 6: Compare suppliers (quality, prices, availability, schedules). This is important in order to avoid undesirable outcomes and to get the best prices. The HB explains how to do this.
Step 7: Come up with a stock-up schedule.

Tips
If you ran out stocks for a product, get used to asking your customer "Do you want something else?". That way, you may find out about demand for products the customer would not have mentioned otherwise.
Reacting to price fluctuations, discounts, and product availability: If price changes are not communicated to you directly by your supplier, you can set a day in the week on which you update those price. Also, keep an open eye on discounts. In our talks with small and mid-sized shop owners, every one of them emphasized this practice as key for success. Sometimes it is your local supermarket which runs a discount offer for cooking oil or a local supplier who needs to get rid of his stock in eggs (and so decreases prices). Do keep an eye on your suppliers' delivery times and stockouts. Do not hesitate to change to a more reliable partner. Retailers like you have emphasized that getting the best prices and being attentive to bulk discounts of their suppliers are keys to reducing stocking-up costs.

Chapter 4: Marketing

Misconception	HB Response
"I have sold the same products for a long time now, there is no need to change my assortment."	In our survey we found that shops that introduced at least one new product in the last three months were associated with significantly higher sales and profits than those that did not.
"If a customer decides to quit buying, I cannot do anything."	From our conversations with shop owners, it became apparent that many do not simply give up on an absent customer, and try to follow up to understand the reasoning.
"I already make the best possible prices, giving discounts is not necessary."	You may want to give discounts for products that are on the verge of spoiling. If your supplier offers you good terms, you can pass on a bit of discount to customers. Also, consider giving discounts for goods bought in bulk.

Reasons to Adopt the Practice
From our survey of small and mid-size retail shops, we gathered that successful shops were the ones that adopted some simple practices to help retain customers and win over new ones.
Shop owners who decide to get back to former customers to see the reasons why they quit buying at their shop associated with monthly sales 24% higher than the sales of businesses whose owners just let it go. What is more, their monthly profits are 23% higher.
From the survey we know that shops that use discounts to attract new and retain loyal customers are associated with monthly sales that are 40% higher than the sales of shops that do not give discounts. Also, their monthly profits are 29% higher.

Chapter 4: Marketing

Step-by-step Implementation Guidance
Step I: Introducing a new product. Though not every customer may be interested in the newly offered products, some may end up buying. There may be customers who do not plan a purchase, but become interested in buying something new after seeing the product in the store.
Step 2: Get back to customers who are no longer buying at your shop. From our survey we know that shop owners who do not give up on former customers and get back in touch have higher sales and profits. It is simple: Almost no customer will mind being asked why they quit buying. So, do not be ashamed to ask. Their reasons can be well-founded. For instance, customers may have moved or may no longer need goods from your shop. Hence, everything is clear. Though, that does not mean you do not have a chance to win back former customers. Another reason can be that someone is simply trying out different vendors, that they happened to pass by other shops, or perhaps prices are better elsewhere. By simply contacting them, you can inquire about their reasons for leaving and also make them feel appreciated. As such, politely asking for their reasons to quit buying at your shop may already win them over. If your customer has been loyal for a long time, on the other hand, you may additionally want to provide a discount to retain their loyalty.
Step 3: Devise a discount strategy. First, it will be useful when new products are introduced. Second, to win back customers you don't see around anymore. Third, for products that are close to expiring. Once you have tried all the different types of discounts, ask yourself which one works best for your shop.

Tips
Even if you do not get a discount from your supplier, they frequently have other special deals. For example, some might leave the product in your shop without you having to pay for it right away. You will only pay later and only for the units you sell. Do not hesitate to ask your supplier for such an offer. Remember, they also want to find out how well the new product sells in the market.
Large suppliers often provide special discounts on new products when those are first introduced to the market. This is to test the product's appeal. Hence, you might want to display the new product in a place that your customers can easily spot. In case your supplier offers you a better price or a special discount for a new product, you may want to pass down the discount to your customers. This way you can see if the item has the potential to sell well or not. Of course, even if you do not receive a better price you can still offer discounts. If the product does not sell well at the reduced price, you know you do not need to sell it again.

Chapter 5: Making Decisions Together

Misconception	HB Response
"I know my shop best, so I do not have to discuss it with anyone."	Nobody doubts that you know your shop best. After all, you have invested your time, money, and energy in it. Plus, you make all the important decisions. Importantly, we do not advise you to give up the decision-making power but to make productive use of other people's inputs, their ideas, and experiences. Hence, this is not about letting others decide in your stead but about making informed decisions incorporating both your knowledge and that of another person. Consider stocking up: Imagine you are in the back of the shop making a list of items to stock up while your spouse is serving customers in front. It can only be beneficial to receive input from your spouse regarding popular items. That way, you will know better what exactly to stock up on and by what amount.
"It will not help me in my business."	Even though you might think discussions and joint decision-making will not help you stocking up or writing records, we have evidence that it benefits your business. In our interviews with small and mid-size retailers all over Jakarta, we found that shop owners who discuss business topics with others are associated with higher sales and profits. In the same way, shop owners who cooperate on any decisions regarding their business also have higher earnings.
"If I discuss my business with people outside of my family, they will use it against me."	There is nothing wrong with starting by consulting your siblings and spouses. In fact, this is a good first step in improving your business. However, do consider seeking out other people to discuss business matters. We understand that not everyone can be trusted but you can still ask for people's opinions without giving away sensitive information. You can explain, in a general fashion, your situation and the decision you face without going into too much detail. You can also start by discussing with your close friends. For business-related advice, you can talk to some of your suppliers. Large suppliers usually have a better overview about what is going on in the market. Surely, you do not want to ask the delivery service whether their products are worth buying – after all, they have an incentive to sell. However, you can ask about other products or new offers. Or simply ask for advice on more general questions such as whether or not your prices are in line with those they see elsewhere.

Chapter 5: Making Decisions Together

Reasons to Adopt the Practice
We found that shop owners who have an open attitude to discussing matters with family members or friends in business or outside of it earn higher profits. Shop owners who said they were open to receiving inputs and who made decisions together with others on any business matter were associated with monthly sales that were 25% higher than the sales of shop owners who did not consult on business matters. In addition, their profits were by 27% higher.
Shop owners who consulted with others on the topic of which new products to offer for sale had 30% higher monthly sales and 35% higher monthly profits compared to shop owners who did not.
Shop owners who decided jointly with others on which new practice to implement in their business made monthly sales 26% higher than the sales of a business whose owner would insist to make decisions alone. In addition, monthly profits were 27% higher.
Two minds can hardly know less than one. This is especially true if the other person differs in their ideas, intuitions, and ways of thinking or if they have expertise different from yours. By including your business partner, co-worker, or spouse in the process of making decisions about the shop, chances are they will become more involved and thus be of even greater use.
It is the most cost-effective way to improve your business. Think about all the things you can do to improve your business. Most come with costs attached or at least some major changes. Consulting others, discussing business matters, and making decisions together is an inexpensive alternative. By exchanging ideas and cooperating with others, you practically benefit for free.

Step-by-step Implementation Guidance
Before we start: Topics to discuss or decide on together. In our survey of 1300 shop owners in Jakarta, we found that many already discuss business matters with their spouses, friends, or suppliers they trust. Most use the opportunity to exchange ideas about sales, selling and purchasing prices, as well as which products are most profitable or sell best. They discuss which products to start offering for sale and where to acquire additional capital (e.g., bank offers or government programs). Moreover, shop owners report discussing discount policies (which discounts to give, to whom, when, and how much), product quality, reliability of suppliers, marketing and promotional offers, and general market trends. Beyond these topics, we strongly recommend discussing business practices.
Step 1: Observe or discuss expertise and define clear roles. The first step in finding out what could be worth discussing apart from the above mentioned topics is to see where you and your co-worker/spouse differ in expertise. Once you know where you differ, you have identified a good discussion topic, you are sure to complement each other's knowledge.
Step 2: Dedicate time for quick feedback. Once you have thought about topics to discuss, whom to discuss them with, and have assigned responsibilities so as to benefit from everyone's expertise, it is a good idea to set out a routine for feedback. It is not easy to break old habits, and so reserving a specific time of the day, preferably towards day's close when you make plans for stocking up, for feedback can help. Just take a few minutes every day to walk over to your friend or to discuss the most important topics in your shop with your co-worker or spouse.
Step 3: Keeping records helps. It may happen that your co-worker or spouse has the feeling that a specific product is selling particularly well without being able to explain. He or she may feel the demand for a product is changing with seasons or in the run-up to holidays but has no way to back it up. Keeping records and updating them whenever sales happen is especially useful here. If both of you keep updating your records, they will both be more comprehensive and will serve as evidence in cases like this. Have a look into your records and see whether the actual sales prove your feeling right.
Step 4: Ask someone outside your family. Finding people outside the family will increase the chance of having more diverse input and different ideas.

Appendix L: Pre-Specified Analysis

This appendix follows the recommendations of Banerjee et al. (2020) and provides a report to clarify any departure of the research analysis from what was pre-registered in the AEA RCT Registry.

This research project was registered in the AEA RCT Registry on 20 April 2016 (access here) under the title, "Constraints to Performance and Growth of Small Retailers in Jakarta," and was assigned the ID number: AEARCTR-0001175.

In what follows, we review each section of the registry entry and clarify any departure of the research analysis from what was pre-registered:

1- Trial Information:

- Pre-registered objectives: The pre-trial states the following three study goals:

"*First, to characterize the business practices used by the retailers, identifying those that are potentially conducive to productivity growth, higher sales, and profits. Second, to disseminate the top performance-enhancing business practices among the retailers and track the adoption of such business practices by retailers. Third, to study the causal effect of adopting the business practices on business performance and growth. In order to distinguish informational and behavioral constraints to adoption of the practices, we vary the ways in which the information is conveyed and implemented*."

These pre-registered objectives comprise the main research questions addressed in the current paper.

The pre-registered study design also states that we will "*make use of different framings as well as business role-models in the implementation of the practices*." Both role-models (in the documentary movie) and frames (in the handbook) are included and described in the current research paper.

- Companion paper**:** A point of clarification is that the pre-registered project feeds into two independent papers: the current (main) paper, "Curating Local Knowledge;" and a companion descriptive paper, "Determinants and Dynamics of Business Aspirations: Evidence from Small-scale Entrepreneurs in an Emerging Market." This companion paper makes use of the rich data on aspirations, preferences, and cognitive attributes that we collected in our surveys to unpack the various forms of business aspirations among retailers in our sample.

2- Sponsors & Partners:

This section is marked as "public" but it appears as "available upon request" due to a programming error at the registry, which thus far remains unresolved. For full information, our pre-registered

sponsors include DFID, Tilburg University, and The World Bank; and our pre-registered implementation partners include J-PAL, SurveyMeter, and Micra.

3- Experimental Details:

- Interventions, Experimental Design, and Experiment Characteristics: The current paper utilizes the same interventions and experimental design that were pre-registered, with two small adjustments. First, the pre-registered sample size was 1,000 retailers, but in practice we were able to increase the sample to 1,040. Second, the pre-registered facilitator intervention included an initial one-hour personal shop visit by a trained facilitator followed by three weekly calls. Based on lessons from our pilot, we implemented a slight variant of this design, with two in-person facilitator visits of 30 minutes each.

- Primary Outcomes: The outcomes pre-registered were business practices, sales, profits, labor productivity, Total Factor Productivity (TFP), cognitive styles, preferences, and aspirations. The current paper includes these outcomes in the analysis with the exception of labor productivity, TFP, and preferences, which we explain below:

    - Labor Productivity: Bloom et al. (2010) defines productivity as average firm-level sales per employee (also labeled labor revenue productivity). Given that the analysis in the main paper does not find treatment effects on the number of employees, we proxy productivity with sales. However, given labor productivity was pre-specified, Appendix Table L1 below presents results for the main treatment specification with labor productivity as the dependent variable.

    - TFP: This variable was originally pre-specified but at the time of survey design and implementation, we recognized the difficulty in measuring TFP, especially with limited time and space in the follow-up surveys. For this reason, we did not measure TFP and do not have results for this outcome.

    - Preferences: We collected data to proxy time and risk preferences with the following two aims: first, data on preferences help better characterize the entrepreneur. Indeed, we use these two measures of preferences to describe the baseline sample (Table 1) and to study selection into treatment (Appendix Table C3). Second, at the time of pre-registration, we thought preferences could change with treatment and that is why we specified them as an outcome. Our hypothesis was based on the literature on endogenous preferences (e.g. Bowles, 1998). However, due to budget constraints we had to cut the length of the endline surveys and we decided not to include these preference questions.

4- Supporting Documents and Materials. This section is private by default and is empty, hence no authorization was made at the time of pre-registration.

5- IRB: No special comment to make. IRB approval was obtained for this project and details are included in the main paper.

6- Analysis Plan (PAP): Our entry in the AEA RCT registry did not include a detailed PAP, which is why the "Analysis Plan" section is empty and appears to be locked. Instead, we pre-registered a description of the project, its objectives, the experiment design, details of the intervention, and the outcomes we would measure. These items have been public since inception.[3]

7- Post-trial Information: This section is empty. The analysis in the main paper and this Appendix provide the full analysis that was pre-specified.

8- Data Publication: The data and replication files will accompany the publication of the paper.

9- Reports, Papers and Other material: This section is empty. No special comment to make.

Appendix L References:

Banerjee, Abhijit, Esther Duflo, Amy Finkelstein, Laurence F. Katz, Benjamin A. Olken, Anja Sautmann. 2020. "In Praise of Moderation: Suggestions for the Scope and Use of Pre-Analysis Plan for RCTs in Economics", *NBER*, w26993.

Bloom, Nicholas, Aprajit Mahajan, David McKenzie and John Roberts. 2010. "Why Do Firms in Developing Countries Have Low Productivity?" *American Economic Review Papers & Proceedings,* 100:2, 619-623.

Bowles, Samuel. 1998. "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions," *Journal of Economic Literature*, American Economic Association, vol. 36(1), pages 75-111, March.

McKenzie, David and Chris Woodruff. 2017. "Business Practices in small firms in developing countries" *Management Science*, 63:9, 2967-2981.

[3] This is consistent with Banerjee et al. (2020), who recommend that the key benefits of a PAP can usually be realized by completing the registration fields in the AEA RCT Registry.

Appendix Table L1: Impact on Pre-specified Outcome -- Labor Productivity

	(1)
	Labor Productivity
Assigned Handbook	-417.378*
	(243.546)
Assigned Handbook & Movie	14.691
	(278.757)
Assigned Handbook & Counseling	171.771
	(266.041)
Assigned All Three	704.530**
	(319.206)
R-squared	0.378
N	2077
Dependent Variable Mean in Control Group	3167.693
Dependent Variable SD in Control Group	4052.670
F-test (p-value): Book = Book & Movie	0.088
F-test (p-value): Book = Book & Counseling	0.014
F-test (p-value): Book = All Three	0.000
F-test (p-value): Book & Movie = Book & Counseling	0.561
F-test (p-value): Book & Movie = All Three	0.039

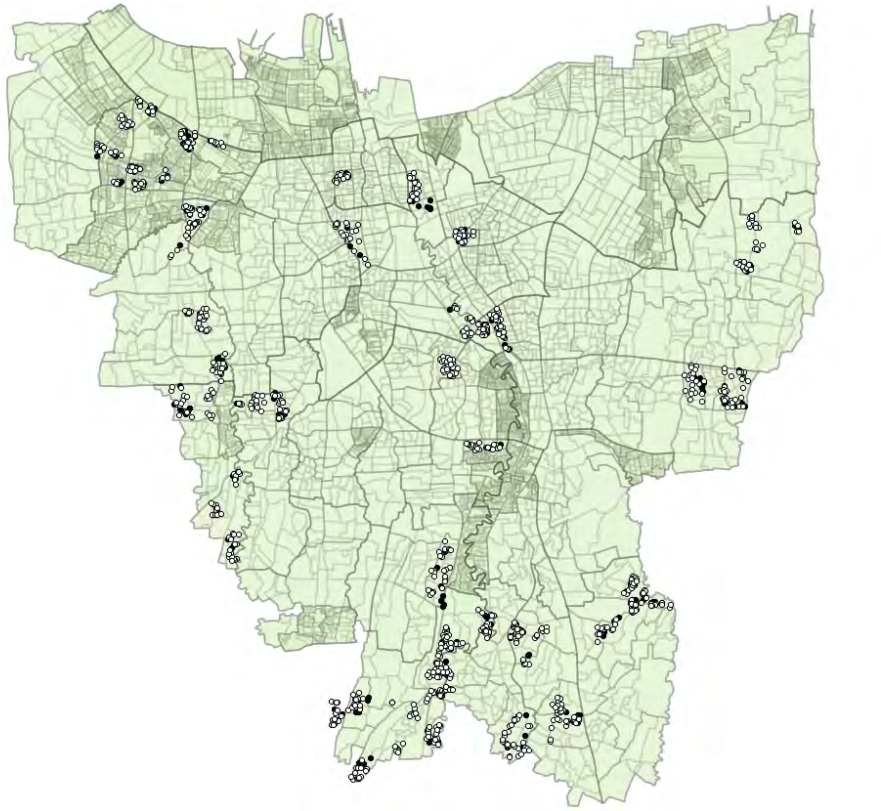
Notes: This table presents regression analysis for total labor productivity, which is defined as the ratio of sales to total number of employees. The sample is pooled across two follow-up survey rounds, a midline at 6 months and endline at 18 months after the intervention. All regression specifications control for a midline survey round dummy, baseline value of the outcome variable, as well as stratification controls. Standard errors in parentheses are clustered at the retail shop level. Statistically significant p-values are highlighted by: * (10% significance level), and ** (5% significance level).

Appendix Figure 1: Businesses Pictures

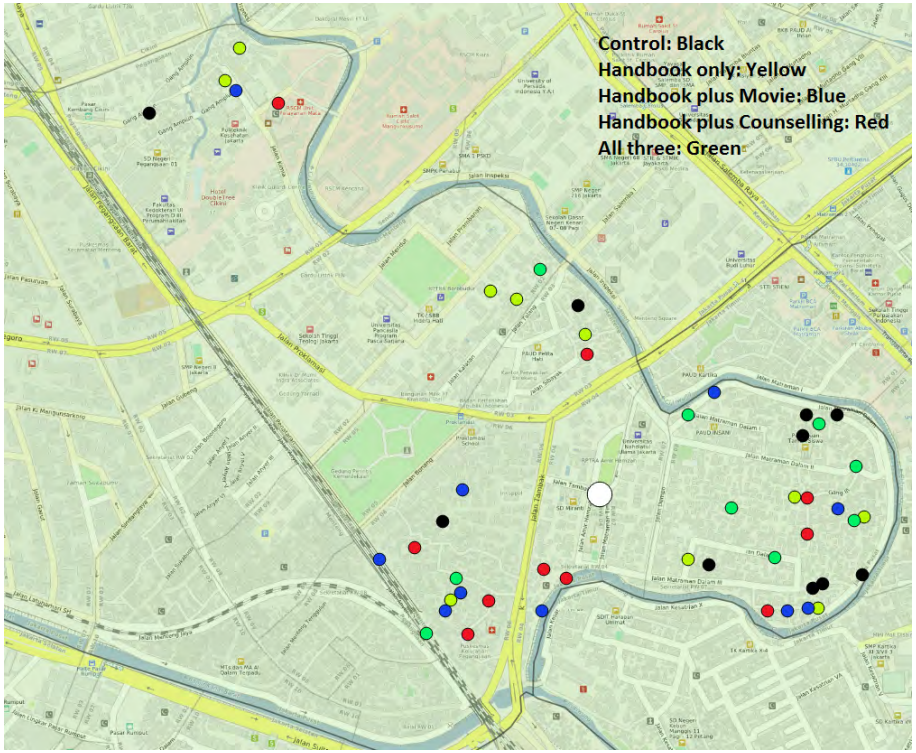
Pictures of two shops representative of the sample of small-scale retail businesses in this study

Appendix Figure 2: Maps of Study Area

(a) Distribution of shop owners in Jakarta (White=Treated; Black=Control)



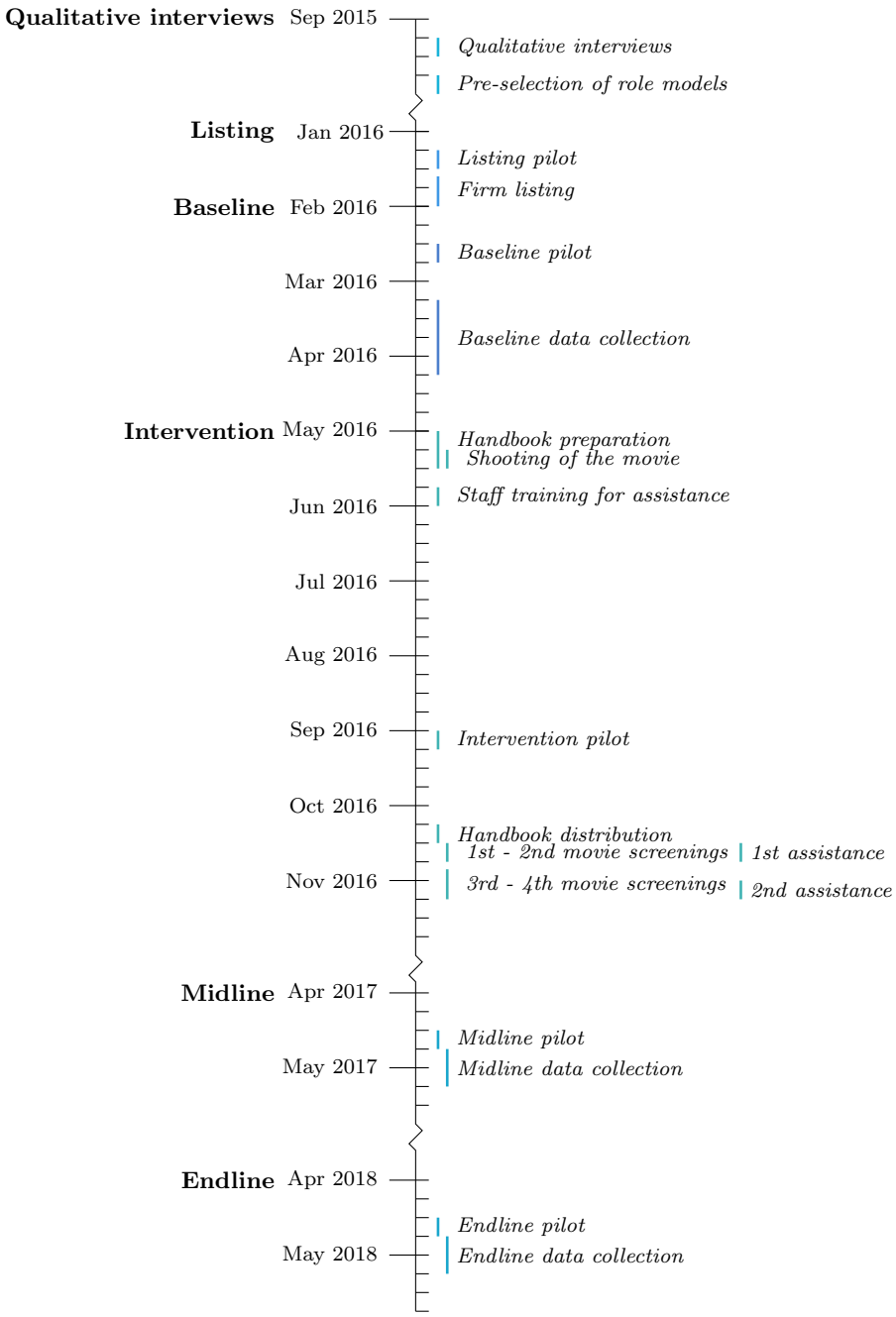
(b) Example Treatment Distribution across shop owners: District Pegangsaan



Appendix Figure 3: Experimental Design

<table>
<tr><td colspan="9">Total Sample
1301 firms</td></tr>
<tr><td rowspan="5">Control
261 firms</td><td colspan="8">Handbooks
1040 firms</td></tr>
<tr><td colspan="8">Returns to Adoption Framing</td></tr>
<tr><td colspan="4">Positive
520 firms</td><td colspan="4">Negative
520 firms</td></tr>
<tr><td colspan="8">Documentary Movie</td></tr>
<tr><td colspan="2">Yes
260 firms</td><td colspan="2">No
260 firms</td><td colspan="2">Yes
260 firms</td><td colspan="2">No
260 firms</td></tr>
<tr><td colspan="8">Assistance</td></tr>
<tr><td>Yes
130
firms</td><td>No
130
firms</td><td>Yes
130
firms</td><td>No
130
firms</td><td>Yes
130
firms</td><td>No
130
firms</td><td>Yes
130
firms</td><td>No
130
firms</td></tr>
</table>

Appendix Figure 4: Project Timeline



Notes: This figure presents the timeline of all field activities for the present study.